```
              GenCore version 5.1.6
        Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on:      April 21, 2005, 04:17:43 ; Search time 219.5 Seconds
                                       (without alignments)
                                       1185.131 Million cell updates/sec

Title:          US-09-016-159D-5
Perfect score:  2715
Sequence:       1 MDHLGASLWPQVGSLCLLLA.........YENSLIPAABPLPPSYVACS 508

Scoring table:  BLOSUM62
                Gapop 10.0 , Gapext 0.5

Searched:       1612378 seqs, 512079187 residues

Total number of hits satisfying chosen parameters:      1612378

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :     UniProt_03:*
               1: uniprot_sprot:*
               2: uniprot_trembl:*

                         SUMMARIES

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.
```

Result No.	Score	% Query Match	Match Length	DB	ID	Description
1	2715	100.0	508	1	EPOR_HUMAN	P19235 homo sapien
2	2283.5	84.1	509	1	EPOR_PIG	Q9myz9 sus scrofa
3	2174.5	80.1	507	1	EPOR_MOUSE	P14753 mus musculu
4	2152.5	79.3	507	1	EPOR_RAT	Q07303 rattus norv
5	1670.5	61.5	387	2	Q95NI3	Q95ni3 ovis aries
6	1645	60.6	418	2	Q95NI4	Q95ni4 ovis aries
7	1058	39.0	316	2	O35545	O35545 rattus norv
8	1016	37.4	229	2	Q27950	Q27950 bos indicus
9	1011	37.2	229	2	Q28206	Q28206 bos taurus
10	529.5	19.5	529	2	Q6UAP7	Q6uap7 tetraodon n
11	251	9.2	522	1	IL9R_HUMAN	Q01113 homo sapien
12	231.5	8.5	630	2	Q6IYE8	Q6iye8 gallus gall
13	224	8.3	566	2	Q925F5	Q925f5 mus musculu
14	220	8.1	625	1	PRLR_PIG	Q6jta8 sus scrofa
15	216.5	8.0	581	2	PRLR_CEREL	Q28235 cervus elap
16	216.5	8.0	625	1	TPOR_MOUSE	O08351 mus musculu
17	215.5	7.9	581	1	PRLR_BOVIN	Q28172 bos taurus
18	215.5	7.9	625	2	Q8BRX0	Q8brx0 mus musculu
19	214.5	7.9	631	2	Q6UAN4	Q6uan4 tetraodon n
20	213	7.8	538	1	I21R_HUMAN	Q9hbe5 homo sapien
21	212.5	7.8	903	2	Q6ICE0	Q6ice0 homo sapien
22	211.5	7.8	635	1	TPOR_HUMAN	P40238 homo sapien
23	210	7.7	529	1	I21R_MOUSE	Q9jhx3 mus musculu
24	210	7.7	616	1	PRLR_RABIT	P14787 oryctolagus
25	208.5	7.7	889	2	Q6NSJ8	Q6nsj8 homo sapien
26	208.5	7.7	897	1	CYRB_HUMAN	P32927 homo sapien
27	207.5	7.6	469	2	Q8C2G1	Q8c2g1 mus musculu
28	206	7.6	467	2	Q63216	Q63216 rattus norv
29	202.5	7.5	581	1	PRLR_SHEEP	O46561 ovis aries
30	202	7.4	468	1	IL9R_MOUSE	Q01114 mus musculu
31	201	7.4	884	2	Q6UAM6	Q6uam6 tetraodon n

32	195	7.2	622	2	Q9NOJ7	Q9n0j7 callithrix
33	192.5	7.1	890	2	Q9Z1A0	Q9z1a0 cavia porce
34	191.5	7.1	610	1	PRLR_RAT	P05710 rattus norv
35	189	7.0	896	2	Q64146	Q64146 rattus sp.
36	189	7.0	896	2	Q78ZF5	Q78zf5 rattus norv
37	188.5	6.9	614	2	Q6EIY6	Q6eiy6 brachydanio
38	187	6.9	819	2	Q6I6F7	Q6i6f7 eublepharis
39	186.5	6.9	896	1	CYRB_MOUSE	P26955 mus musculu
40	184.5	6.8	622	1	PRLR_HUMAN	P16471 homo sapien
41	184	6.8	831	1	PRLR_CHICK	Q04594 gallus gall
42	183.5	6.8	830	1	PRLR_COLLI	Q90374 columba liv
43	182.5	6.7	831	2	Q6QDAO	Q6qda0 gallus gall
44	182	6.7	896	2	Q8QZX9	Q8qzx9 mus musculu
45	180	6.6	608	1	PRLR_MOUSE	Q08501 mus musculu

```
                         ALIGNMENTS

RESULT 1
EPOR_HUMAN
ID   EPOR_HUMAN       STANDARD;      PRT;     508 AA.
AC   P19235; Q15443;
DT   01-NOV-1990 (Rel. 16, Created)
DT   01-NOV-1990 (Rel. 16, Last sequence update)
DT   25-OCT-2004 (Rel. 45, Last annotation update)
DE   Erythropoietin receptor precursor (EPO-R).
GN   Name=EPOR;
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A. (ISOFORM EPOR-F).
RC   TISSUE=Erythroleukemia, and Fetal liver;
RX   MEDLINE=90304340; PubMed=2163696;
RA   Jones S.S., D'Andrea A.D., Haines L.L., Wong G.G.;
RT   "Human erythropoietin receptor: cloning, expression, and biologic
RT   characterization.";
RL   Blood 76:31-35(1990).
RN   [2]
RP   SEQUENCE FROM N.A. (ISOFORM EPOR-F).
RC   TISSUE=Placenta;
RX   MEDLINE=92399733; PubMed=1668606;
RA   Noguchi C.T., Bae K.S., Chin K., Wada Y., Schechter A.N.,
RA   Hankins W.D.;
RT   "Cloning of the human erythropoietin receptor gene.";
RL   Blood 78:2548-2556(1991).
RN   [3]
RP   SEQUENCE FROM N.A. (ISOFORM EPOR-F).
RC   TISSUE=Erythroleukemia;
RX   MEDLINE=91372359; PubMed=1654273;
RA   Ehrenman K., St John T.;
RT   "The erythropoietin receptor gene: cloning and identification of
RT   multiple transcripts in an erythroid cell line OCIM1.";
RL   Exp. Hematol. 19:973-977(1991).
RN   [4]
RP   SEQUENCE FROM N.A. (ISOFORMS EPOR-F; EPOR-S AND EPOR-T).
RC   TISSUE=Bone marrow, and Megakaryoblast;
RX   PubMed=1324524;
RA   Nakamura Y., Komatsu N., Nakauchi H.;
RT   "A truncated erythropoietin receptor that fails to prevent programmed
RT   cell death of erythroid cells.";
RL   Science 257:1138-1141(1992).
RN   [5]
RP   SEQUENCE OF 1-96 FROM N.A.
RC   TISSUE=Placenta;
RX   MEDLINE=92399734; PubMed=1668607;
RA   Maouche L., Tournamille C., Hattab C., Boffa G., Cartron J.-P.,
RA   Chretien S.;
RT   "Cloning of the gene encoding the human erythropoietin receptor.";
RL   Blood 78:2557-2563(1991).
RN   [6]
```

RP   SEQUENCE OF 1-17 FROM N.A.
RX   MEDLINE=92147143; PubMed=1664413;
RA   Penny L.A., Forget B.G.;
RT   "Genomic organization of the human erythropoietin receptor gene.";
RL   Genomics 11:974-980(1991).
RN   [7]
RP   PARTIAL SEQUENCE FROM N.A. (EPOR-S).
RC   TISSUE=Erythroid cells;
RX   MEDLINE=92039091; PubMed=1657727; DOI=10.1016/0378-1119(91)90213-U;
RA   Todokoro K., Kuramochi S., Nagasawa T., Abe T., Ikawa Y.;
RT   "Isolation of a cDNA encoding a potential soluble receptor for human
RT   erythropoietin.";
RL   Gene 106:283-284(1991).
RN   [8]
RP   STAT1/STAT3 ACTIVATION, AND MUTAGENESIS OF TYR-456 AND TYR-468.
RX   PubMed=11756159; DOI=10.1182/blood.V99.1.102;
RA   Kirito K., Nakajima K., Watanabe T., Uchida M., Tanaka M., Ozawa K.,
RA   Komatsu N.;
RT   "Identification of the human erythropoietin receptor region required
RT   for Stat1 and Stat3 activation.";
RL   Blood 99:102-110(2002).
RN   [9]
RP   INTERACTION WITH PTPN11.
RX   PubMed=7534299; DOI=10.1074/jbc.270.10.5631;
RA   Tauchi T., Feng G.-S., Shen R., Hoatlin M., Bagby G.C. Jr., Kabat D.,
RA   Lu L., Broxmeyer H.E.;
RT   "Involvement of SH2-containing phosphotyrosine phosphatase Syp in
RT   erythropoietin receptor signal transduction pathways.";
RL   J. Biol. Chem. 270:5631-5635(1995).
RN   [10]
RP   LIGAND BINDING SITE, AND MUTAGENESIS OF THR-114; SER-115; SER-116;
RP   PHE-117; VAL-118; LEU-120; GLU-121; ARG-165; MET-174; SER-176; HIS-177
RP   AND ARG-179.
RX   PubMed=8662939; DOI=10.1074/jbc.271.24.14045;
RA   Middleton S.A., Johnson D.L., Jin R., McMahon P.J., Collins A.,
RA   Tullai J., Gruninger R.H., Jolliffe L.K., Mulcahy L.S.;
RT   "Identification of a critical ligand binding determinant of the human
RT   erythropoietin receptor. Evidence for common ligand binding motifs in
RT   the cytokine receptor family.";
RL   J. Biol. Chem. 271:14045-14054(1996).
RN   [11]
RP   SUBCELLULAR LOCATION OF ISOFORM EPOR-S.
RX   PubMed=11722595;
RA   Motohashi T., Nakamura Y., Osawa M., Hiroyama T., Iwama A.,
RA   Shibuya A., Nakauchi H.;
RT   "Increased cell surface expression of C-terminal truncated
RT   erythropoietin receptors in polycythemia.";
RL   Eur. J. Haematol. 67:88-93(2001).
RN   [12]
RP   INTERACTION WITH SOCS3, AND MUTAGENESIS OF TYR-454 AND TYR-456.
RX   MEDLINE=22024229; PubMed=12027890;
RA   Hoertner M., Nielsch U., Mayr L.M., Heinrich P.C., Haan S.;
RT   "A new high affinity binding site for suppressor of cytokine
RT   signaling-3 on the erythropoietin receptor.";
RL   Eur. J. Biochem. 269:2516-2526(2002).
RN   [13]
RP   INTERACTION WITH NOSIP.
RX   PubMed=12746455; DOI=10.1074/jbc.M210039200;
RA   Friedman A.D., Nimbalkar D., Quelle F.W.;
RT   "Erythropoietin receptors associate with a ubiquitin ligase, p33RUL,
RT   and require its activity for erythropoietin-induced proliferation.";
RL   J. Biol. Chem. 278:26851-26861(2003).
RN   [14]
RP   PHOSPHORYLATION, AND INTERACTION WITH APS.
RX   MEDLINE=99301417; PubMed=10374881;
RA   Wakioka T., Sasaki A., Mitsui K., Yokouchi M., Inoue A., Komiya S.,
RA   Yoshimura A.;
RT   "APS, an adaptor protein containing Pleckstrin homology (PH) and Src
RT   homology-2 (SH2) domains inhibits the JAK-STAT pathway in
RT   collaboration with c-Cbl.";
RL   Leukemia 13:760-767(1999).
RN   [15]
RP   INTERACTION WITH ATXN2L.

RX   MEDLINE=21883980; PubMed=11784712; DOI=10.1074/jbc.M105970200;
RA   Meunier C.F., Bordereaux D., Porteu F., Gisselbrecht S., Chretien S.,
RA   Courtois G.;
RT   "Cloning and characterization of a family of proteins associated with
RT   Mpl.";
RL   J. Biol. Chem. 277:9139-9147(2002).
RN   [16]
RP   X-RAY CRYSTALLOGRAPHY (2.8 ANGSTROMS) OF 34-244.
RX   MEDLINE=96291992; PubMed=862530;
RA   Livnah O., Stura E.A., Johnson D.L., Middleton S.A., Mulcahy L.S.,
RA   Wrighton N.C., Dower W.J., Jolliffe L.K., Wilson I.A.;
RT   "Functional mimicry of a protein hormone by a peptide agonist: the EPO
RT   receptor complex at 2.8 A.";
RL   Science 273:464-471(1996).
RN   [17]
RP   X-RAY CRYSTALLOGRAPHY (2.7 ANGSTROMS) OF 34-244.
RX   MEDLINE=99023198; PubMed=9808045;
RA   Livnah O., Johnson D.L., Stura E.A., Farrell F.X., Barbone F.P.,
RA   You Y., Liu K.D., Goldsmith M.A., He W., Krause C.D., Pestka S.,
RA   Jolliffe L.K., Wilson I.A.;
RT   "An antagonist peptide-EPO receptor complex suggests that receptor
RT   dimerization is not sufficient for activation.";
RL   Nat. Struct. Biol. 5:993-1004(1998).
RN   [18]
RP   X-RAY CRYSTALLOGRAPHY (1.9 ANGSTROMS) OF 32-244 IN COMPLEX WITH EPO.
RX   MEDLINE=98445092; PubMed=9774108; DOI=10.1038/26773;
RA   Syed R.S., Reid S.W., Li C., Cheetham J.C., Aoki K.H., Liu B.,
RA   Zhan H., Osslund T.D., Chirino A.J., Zhang J., Finer-Moore J.,
RA   Elliott S., Sitney K., Katz B.A., Matthews D.J., Wendoloski J.J.,
RA   Egrie J., Stroud R.M.;
RT   "Efficiency of signalling through cytokine receptors depends
RT   critically on receptor orientation.";
RL   Nature 395:511-516(1998).
RN   [19]
RP   X-RAY CRYSTALLOGRAPHY (2.40 ANGSTROMS) OF 34-246.
RX   MEDLINE=99141272; PubMed=9974392; DOI=10.1126/science.283.5404.987;
RA   Livnah O., Stura E.A., Middleton S.A., Johnson D.L., Jolliffe L.K.,
RA   Wilson I.A.;
RT   "Crystallographic evidence for preformed dimers of erythropoietin
RT   receptor before ligand activation.";
RL   Science 283:987-990(1999).
RN   [20]
RP   INVOLVEMENT IN FAMILIAL ERYTHROCYTOSIS.
RX   PubMed=8506290;
RA   de la Chapelle A., Traskelin A.-L., Juvonen E.;
RT   "Truncated erythropoietin receptor causes dominantly inherited benign
RT   human erythrocytosis.";
RL   Proc. Natl. Acad. Sci. U.S.A. 90:4495-4499(1993).
-!- FUNCTION: Receptor for erythropoietin. Mediates erythropoietin-
CC   induced erythroblast proliferation and differentiation. Upon EPO
CC   stimulation, EPOR dimerizes triggering the JAK2/STAT5 signaling
CC   cascade. In some cell types, can also activate STAT1 and STAT3.
CC   May also activate the LYN tyrosine kinase.
-!- FUNCTION: Isoform EPOR-T, missing the cytoplasmic tail, acts as a
CC   dominant-negative receptor of EPO-mediated signaling.
-!- SUBUNIT: Forms homodimers on EPO stimulation. The tyrosine-
CC   phosphorylated form interacts with several SH2 domain-containing
CC   proteins including LYN (By similarity), the adapter protein APS,
CC   PTPN6 (By similarity), PTPN11, JAK2, PI3 kinases, STAT5A/B, SOCS3,
CC   CRKL (By similarity). The N-terminal SH2 domain of PTPN6 binds
CC   Tyr-454 and inhibits signaling through dephosphorylation of JAK2
CC   (By similarity). APS binding also inhibits the JAK-STAT signaling.
CC   Binding to PTPN11, preferentially through the N-terminal SH2
CC   domain, promotes mitogenesis and phosphorylation of PTPN11 (By
CC   similarity). Binding of JAK2 (through its N-terminal) promotes
CC   cell-surface expression (By similarity). Interaction with the
CC   ubiquitin ligase NOSIP mediates EPO-induced cell proliferation.
CC   Interacts with ATXN2L.
-!- SUBCELLULAR LOCATION: Type I membrane protein. Isoform EPOR-S is
CC   secreted and located to the cell surface.
-!- ALTERNATIVE PRODUCTS:
CC   Event=Alternative splicing; Named isoforms=3;
CC   Name=EPOR-F; Synonyms=Full-length form;

CC   IsoId=P19235-1; Sequence=Displayed;
CC   Name=EPOR-S; Synonyms=Soluble form;

Query Match   100.0%;   Score 2715;   DB 1;   Length 508;
Best Local Similarity   100.0%;   Pred. No. 4.8e-181;
Matches 508; Conservative   0; Mismatches   0; Indels   0; Gaps   0;

Qy   1 MDHLGASLWPQVGSLCLLLAGAAWAPPNLPDPKFESKAALLAARGPEELLCFTERLEDL 60
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1 MDHLGASLWPQVGSLCLLLAGAAWAPPNLPDPKFESKAALLAARGPEELLCFTERLEDL 60

Qy   61 VCFWEEAASAGVGPGNYSFSYQLEDEPWKLCRLHQAPTARGAVRFWCSLPTADTSSFVPL 120
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61 VCFWEEAASAGVGPGNYSFSYQLEDEPWKLCRLHQAPTARGAVRFWCSLPTADTSSFVPL 120

Qy 121 ELRVTAASGAPRYHRVIHINEVVLLDAPVGLVARLADESGHVVLRWLPPPETPMTSHIRY 180
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 121 ELRVTAASGAPRYHRVIHINEVVLLDAPVGLVARLADESGHVVLRWLPPPETPMTSHIRY 180

Qy 181 EVDVSAGNGAGSVQRVEILEGRTECVLSNLRGRTRYTFAVRARMAEPSFGGFWSAWSEPV 240
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 181 EVDVSAGNGAGSVQRVEILEGRTECVLSNLRGRTRYTFAVRARMAEPSFGGFWSAWSEPV 240

Qy 241 SLLTPSDLDPLILTLSLILVVILVLLTVLALLSHRRALKQKIWPGIPSPESEFEGLFTTH 300
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 241 SLLTPSDLDPLILTLSLILVVILVLLTVLALLSHRRALKQKIWPGIPSPESEFEGLFTTH 300

Qy 301 KGNFQLWLYQNDGCLWWSPCTPFTEDPPASLEVLSERCWGTMQAVEPGTDDEGPLLEPVG 360
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 301 KGNFQLWLYQNDGCLWWSPCTPFTEDPPASLEVLSERCWGTMQAVEPGTDDEGPLLEPVG 360

Qy 361 SEHAQDTYLVLDKWLLPRNPPSEDLPGPGGSVDIVAMDEGSEASSCSSALASKPSPEGAS 420
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 361 SEHAQDTYLVLDKWLLPRNPPSEDLPGPGGSVDIVAMDEGSEASSCSSALASKPSPEGAS 420

Qy 421 AASFEYTILDPSSQLLRPWTLCPELPPTPHLKYLYLVVSDGISTDYSSGDSQGAQGGL 480
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 421 AASFEYTILDPSSQLLRPWTLCPELPPTPHLKYLYLVVSDGISTDYSSGDSQGAQGGL 480

Qy 481 SDGPYSNPYENSLIPAAEPLPPSYVACS 508
         ||||||||||||||||||||||||||||
Db 481 SDGPYSNPYENSLIPAAEPLPPSYVACS 508

RESULT 2
EPOR_PIG
ID   EPOR_PIG   STANDARD;   PRT;   509 AA.
AC   Q9MYZ9;
DT   29-MAR-2004 (Rel. 43, Created)
DT   29-MAR-2004 (Rel. 43, Last sequence update)
DT   25-OCT-2004 (Rel. 45, Last annotation update)
DE   Erythropoietin receptor precursor (EPO-R).
GN   Name=EPOR;
OS   Sus scrofa (Pig).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
OX   NCBI_TaxID=9823;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Liver;
RX   PubMed=10962196; DOI=10.1016/S0739-7240(00)00062-X;
RA   Pearson P.L., Smith T.P.L., Sonstegard T.S., Klemcke H.G.,
RA   Christenson R.K., Vallet J.L.;
RT   "Porcine erythropoietin receptor: molecular cloning and expression in
RT   embryonic and fetal liver.";
RL   Domest. Anim. Endocrinol. 19:25-38(2000).
CC   -!- FUNCTION: Receptor for erythropoietin. Mediates erythropoietin-
CC   induced erythroblast proliferation and differentiation. Upon EPO
CC   stimulation, EPOR dimerizes triggering the JAK2/STATS signaling
CC   cascade. In some cell types, can also activate STAT1 and STAT3.
CC   May also activate LYN tyrosine kinase (By similarity).
CC   -!- SUBUNIT: Forms homodimers on EPO stimulation. The tyrosine-
CC   phosphorylated form interacts with several SH2 domain-containing

CC   proteins including LYN, the adapter protein APS, PTPN6, PTPN11,
CC   JAK2, PI3 kinases, STAT5A/B, SOCS3 and CRKL. The N-terminal SH2
CC   domain of PTPN6 binds Tyr-455 and inhibits signaling through
CC   dephosphorylation of JAK2. APS binding also inhibits the JAK-STAT
CC   signaling. Binding to PTPN1, preferentially through the N-
CC   terminal SH2 domain, promotes mitogenesis and phosphorylation of
CC   PTPN11. Binding of JAK2 (through its N-terminal) promotes cell-
CC   surface expression. Interaction with the ubiquitin ligase NOSIP
CC   mediates EPO-induced cell proliferation. Interacts with ATXN2L (By
CC   similarity).
CC   -!- SUBCELLULAR LOCATION: Type I membrane protein.
CC   -!- DEVELOPMENTAL STAGE: Low expression at day 24 gestation in fetal
CC   liver. Expression increases dramatically thereafter to day 30.
CC   Levels then remain constant up to day 40.
CC   -!- DOMAIN: The WSXWS motif appears to be necessary for proper protein
CC   folding and thereby efficient intracellular transport and cell-
CC   surface receptor binding.
CC   -!- DOMAIN: The box 1 motif is required for JAK interaction and/or
CC   activation.
CC   -!- DOMAIN: Contains 1 copy of a cytoplasmic motif that is referred to
CC   as the immunoreceptor tyrosine-based inhibitor motif (ITIM). This
CC   motif is involved in modulation of cellular responses. The
CC   phosphorylated ITIM motif can bind the SH2 domain of several SH2-
CC   containing phosphatases.
CC   -!- PTM: On EPO stimulation, phosphorylated on C-terminal tyrosine
CC   residues by JAK2. The phosphotyrosine motifs are also docking
CC   sites for several SH2-containing proteins and adapter proteins
CC   which mediate cell proliferation. Phosphorylation on Tyr-455 is
CC   required for PTPN6 interaction, Tyr-427 for PTPN11. Tyr-427 is
CC   also required for SOCS3 binding, but Tyr-455/Tyr-457 motif is the
CC   preferred binding site (By similarity).
CC   -!- PTM: Ubiquitinated by NOSIP; appears to be either multi-
CC   monoubiquitinated or polyubiquitinated. Ubiquitination mediates
CC   proliferation and survival of EPO-dependent cells (By similarity).
CC   -!- SIMILARITY: Belongs to the type I cytokine family of receptors.
CC   Subfamily 1.
CC   -!- SIMILARITY: Contains 1 fibronectin type III domain.

DR   EMBL; AF274305; AAF77065.1; -.
DR   HSSP; P19235; 1ERN.
DR   InterPro; IPR002996; Cytkn_recept_B/G.
DR   InterPro; IPR009167; EPO_receptor.
DR   InterPro; IPR003961; FN_III.
DR   InterPro; IPR008957; FN_III-like.
DR   InterPro; IPR003528; Hemptrecept_Fl.
DR   Pfam; PF00041; fn3; 1.
DR   PIRSF; PIRSF001959; EPO_receptor; 1.
DR   SMART; SM00060; FN3; 1.
DR   PROSITE; PS50853; FN3; 1.
DR   PROSITE; PS01352; HEMATOPO_REC_L_F1; 1.
KW   Glycoprotein; Phosphorylation; Receptor; Signal; Transmembrane;
KW   Ubl conjugation.
FT   SIGNAL   1   24   Potential.
FT   CHAIN   25   509   Erythropoietin receptor.
FT   DOMAIN   25   251   Extracellular (Potential).
FT   TRANSMEM   252   274   Potential.
FT   DOMAIN   275   509   Cytoplasmic (Potential).
FT   DOMAIN   145   241   Fibronectin type-III.
FT   SITE   117   117   Required for ligand binding (By
FT   similarity).
FT   SITE   234   238   WSXWS motif.
FT   SITE   282   291   Box 1 motif.
FT   SITE   369   369   APS binding, and STAT5 binding and
FT   activation (By similarity).
FT   SITE   427   427   Required for STAT5/PTPN11/SOCS3 binding.

```
FT   SITE            453    458       ITIM motif.
FT   SITE            455              PTPN6 binding.
FT   DISULFID        52     62        By similarity.
FT   DISULFID        91     107       By similarity.
FT   MOD_RES         369              Phosphotyrosine (by JAK2) (By
FT                                    similarity).
FT   MOD_RES         427              Phosphotyrosine (by JAK2) (By
FT                                    similarity).
FT   MOD_RES         455              Phosphotyrosine (by JAK2) (By
FT                                    similarity).
FT   MOD_RES         457              Phosphotyrosine (by JAK2) (By
FT                                    similarity).
FT   MOD_RES         469              Phosphotyrosine (by JAK2) (By
FT                                    similarity).
FT   MOD_RES         486              Phosphotyrosine (by JAK2) (By
FT                                    similarity).
FT   MOD_RES         490              Phosphotyrosine (by JAK2) (By
FT                                    similarity).
FT   MOD_RES         505              Phosphotyrosine (by JAK2) (By
FT                                    similarity).
FT   CARBOHYD        184    184       N-linked (GlcNAc...) (Potential).
SQ   SEQUENCE   509 AA;  55183 MW;  35B565D07C6BCD8A CRC64;

Query Match            84.1%;   Score 2283.5;  DB 1;  Length 509;
Best Local Similarity  84.5%;   Pred. No. 6e-151;  Indels 1;  Gaps 1;
Matches 430;  Conservative 17;  Mismatches 17;

Qy    1   MDHLGASLWPQVGSLCLLLAGAAWAPPNLPDPKFESKAALLAARGPEELLCFTERLEDL   60
              |   || ||||||||||||||| || ||  || |||||||||||||||||||||||||||
Db    1   MYHFGATLWPGVGSLCLLLAGATWAPSPNSPDAKFESKAALLAARGPEELLCFTERLEDL   60

Qy   61   VCFWEEAASAGVGPGNYSFSYQLEDEPWKLCRLHQAPTARGAVRFWCSLPTADTSSFVPL  120
              ||||||  |||||| |  ||||||||| |   || || |||||| ||||||||||||||||||
Db   61   VCFWEEAGSAGVGPEDYSFSYQLEGEPWKPCHLHQGPTARGSVRFWCSLPTADTSSFVPL  120

Qy  121   ELRVT-AASGAPRYHRVIHINEVVLLDAPVGLVARLADESGHVVLRWLPPPETPMTSHIR  179
              ||||| | ||||||||| ||||||||| ||| ||  | |||||||||||||   || |||
Db  121   ELRVTEVSSGAPRYHRIIHINEVVLLDPPAGLLARRAEESGHVVLRWLPPPGAPMASLIR  180

Qy  180   YEVDVSAGNGAGSVQRVEILEGRTECVLSNLRGRTRYTFAVRARMAEPSFGGFWSAWSEP  239
              |||   |  ||| |||||||| ||||||||||| |||||| ||||||||||||||||||||
Db  181   YEVNISTENAAGVQRVEILDGRTECVLSNLRGGTRYTFMVRARMAEPSFGGFWSAWSEP   240

Qy  240   VSLLTPSDLDPLILTLSLILVVILVLLTVLALLSHRRALKQKIWPGIPSPESEFEGLFTT  299
               ||||  ||||||||||||||| | ||| ||||||||| |||||||||||||| ||||||||
Db  241   ASLLTASDLDPLILTLSLILVILILLLAVLALLSHRRTLKQKIWPGIPSPEGEFEGLFTT  300

Qy  300   HKGNFQLWLYQNDGCLWWSPCTPFTEDPPASLEVLSERCWGTMQAVEPGTDEGPLLEPV   359
              ||||||||||| ||||||||||||| ||  |  ||||||||||   ||||||    |||| ||
Db  301   HKGNFQLWLYQTDGCLWWSPCTPFAEDDPAPLEVLSERCWGVTQAVEPAADDEGSLLEPV  360

Qy  360   GSEHAQDTYLVLDKWLLPRNPPSEDLPGPGGSVDIVAMDEGSEASSCSSALASKPSPEGA  419
              |||| ||||||||||||||  | |||| | |   ||  || || ||  ||||| | |||||
Db  361   GSEHARDTYLVLDKWLLPRRPASEDLPQPGGPGGDLDWAAMDEASEASFCSSALALKPGPEGA  420

Qy  420   SAASFEYTILDPSSQLLRPWTLCPELPPTPPHLKYLYLVVSDGISTDYSSGDSQGAQGG  479
              ||||||||||||||||||||     |  ||||||||||||||||||||||||||  ||  ||||
Db  421   SAASFEYTILDPSSQLLRPRALPAELPPTPPHLKYLYLVVSDGISTDYSSGGSQETQGG  480

Qy  480   LSDGPYSNPYENSLIPAAEPLPPSYVACS  508
               | ||||||||||||| | ||  || || |
Db  481   SSSGPYSNPYENSLVPAPEPSPPNYVVCS  509


RESULT 3
EPOR_MOUSE
ID   EPOR_MOUSE      STANDARD;      PRT;      507 AA.
AC   P14753; Q63852;
DT   01-APR-1990 (Rel. 14, Created)
DT   01-APR-1990 (Rel. 14, Last sequence update)
DT   25-OCT-2004 (Rel. 45, Last annotation update)
DE   Erythropoietin receptor precursor (EPO-R).
GN   Name=Epor;
```

```
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A. (ISOFORM EPOR-F).
RC   TISSUE=Erythroleukemia;
RX   MEDLINE=89195238; PubMed=2539263; DOI=10.1016/0092-8674(89)90965-3;
RA   D'Andrea A.D., Lodish H.F., Wong G.G.;
RT   "Expression cloning of the murine erythropoietin receptor.";
RL   Cell 57:277-285(1989).
RN   [2]
RP   SEQUENCE FROM N.A. (ISOFORMS EPOR-F AND EPOR-S).
RC   STRAIN=BALB/c; TISSUE=Erythroleukemia, and Liver;
RX   MEDLINE=91080149; PubMed=2175360;
RA   Kuramochi S., Ikawa Y., Todokoro K.;
RT   "Characterization of murine erythropoietin receptor genes.";
RL   J. Mol. Biol. 216:567-575(1990).
RN   [3]
RP   SEQUENCE FROM N.A. (ISOFORM EPOR-F), AND ALTERNATIVE SPLICING DUE TO
RP   FRIEND SPLEEN FOCUS-FORMING VIRUS.
RC   TISSUE=Erythroleukemia;
RX   MEDLINE=92017832; PubMed=1656233;
RA   Hino M., Tojo A., Misawa Y., Morii H., Takaku F., Shibuya M.;
RT   "Unregulated expression of the erythropoietin receptor gene caused by
RT   insertion of spleen focus-forming virus long terminal repeat in a
RT   murine erythroleukemia cell line.";
RL   Mol. Cell. Biol. 11:5527-5533(1991).
RN   [4]
RP   SEQUENCE FROM N.A. (ISOFORM EPOR-F).
RC   STRAIN=C57BL/6J, and FVB/N-3; TISSUE=Breast tumor;
RX   MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA   Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA   Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA   Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA   Hopkins R.F., Jordan H.F., Moore T., Max S.I., Wang J., Hsieh F.,
RA   Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA   Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA   Brownstein M.J., Usdin T.B., Toshiyuki C., Carninci P., Prange C.,
RA   Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA   Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA   Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA   Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA   Fahey J., Helton E., Ketteman M., Madan A., Rodrigues S., Sanchez A.,
RA   Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA   Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA   Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA   Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
RA   Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT   "Generation and initial analysis of more than 15,000 full-length human
RT   and mouse cDNA sequences.";
RL   Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN   [5]
RP   SEQUENCE OF 1-27 FROM N.A.
RX   MEDLINE=90287158; PubMed=2162479;
RA   Youssoufian H., Zon L.I., Orkin S.H., D'Andrea A.D., Lodish H.F.;
RT   "Structure and transcription of the mouse erythropoietin receptor
RT   gene.";
RL   Mol. Cell. Biol. 10:3675-3682(1990).
RN   [6]
RP   SEQUENCE OF 1-24 FROM N.A.
RX   MEDLINE=91201346; PubMed=1849897;
RA   Lacombe C., Chretien S., Lemarchandel V., Mayeux P., Romeo P.H.,
RA   Gisselbrecht S., Cartron J.-P.;
RT   "Spleen focus-forming virus long terminal repeat insertional
RT   activation of the murine erythropoietin receptor gene in the T3Cl-2
RT   friend leukemia cell line.";
RL   J. Biol. Chem. 266:6952-6956(1991).
RN   [7]
RP   FUNCTION IN ERYTHROPOIETIN-INDUCED MITOGENESIS, AND MUTAGENESIS OF
RP   GLN-304, TRP-306, SER-317, GLU-324, LEU-330 AND GLU-331.
RX   MEDLINE=93180826; PubMed=8382775;
RA   Miura O., Cleveland J.L., Ihle J.N.;
```

RT   "Inactivation of erythropoietin receptor function by point mutations
RT   in a region having homology with other cytokine receptors.";
RL   Mol. Cell. Biol. 13:1788-1795(1993).
RN   [8]
RP   FUNCTION OF THE WSXWS MOTIF, AND MUTAGENESIS OF TRP-232; SER-233;
RP   ALA-234; TRP-235 AND SER-236.
RX   PubMed=8617735; DOI=10.1074/jbc.271.9.4699;
RA   Hilton D.J., Watowich S.S., Katz L., Lodish H.F.;
RT   "Saturation mutagenesis of the WSXWS motif of the erythropoietin
RT   receptor.";
RL   J. Biol. Chem. 271:4699-4708(1996).
RN   [9]
RP   INTERACTION WITH JAK2, PHOSPHORYLATION, AND MUTAGENESIS OF TRP-306.
RX   PubMed=8068943;
RA   Miura O., Nakamura N., Quelle F.W., Witthuhn B.A., Ihle J.N., Aoki N.;
RT   "Erythropoietin induces association of the JAK2 protein tyrosine
RT   kinase with the erythropoietin receptor in vivo.";
RL   Blood 84:1501-1507(1994).
RN   [10]
RP   INTERACTION WITH PTPN6, AND MUTAGENESIS OF TYR-453; 453-TYR--TYR-455
RP   AND TYR-455.
RX   PubMed=7889566; DOI=10.1016/0092-8674(95)90351-8;
RA   Klingmueller U., Lorenz U., Cantley L.C., Neel B.G., Lodish H.F.;
RT   "Specific recruitment of SH-PTP1 to the erythropoietin receptor causes
RT   inactivation of JAK2 and termination of proliferative signals.";
RL   Cell 80:729-738(1995).
RN   [11]
RP   INTERACTION WITH PTPN11, AND MUTAGENESIS OF TYR-367; TYR-425; TYR-453;
RP   TYR-455; TYR-467; TYR-484; TYR-488 AND TYR-503.
RX   PubMed=8639815;
RA   Tauchi T., Damen J.E., Toyama K., Feng G.-S., Broxmeyer H.E.,
RA   Krystal G.;
RT   "Tyrosine 425 within the activated erythropoietin receptor binds Syp,
RT   reduces the erythropoietin required for Syp tyrosine phosphorylation,
RT   and promotes mitogenesis.";
RL   Blood 87:4495-4501(1996).
RN   [12]
RP   INTERACTION WITH STAT5.
RX   PubMed=8665851;
RA   Gobert S., Chretien S., Gouilleux F., Muller O., Pallard C.,
RA   Dusanter-Fourt I., Groner B., Lacombe C., Gisselbrecht S., Mayeux P.;
RT   "Identification of tyrosine residues within the intracellular domain
RT   of the erythropoietin receptor crucial for STAT5 activation.";
RL   EMBO J. 15:2434-2441(1996).
RN   [13]
RP   INTERACTION WITH STAT5, AND MUTAGENESIS OF ARG-153; GLN-304; TRP-306;
RP   SER-317; GLU-324; LEU-330; GLU-331 AND TYR-367.
RX   PubMed=8657137;
RA   Quelle F.W., Wang D., Nosaka T., Thierfelder W.E., Stravopodis D.,
RA   Weinstein Y., Ihle J.N.;
RT   "Erythropoietin induces activation of Stat5 through association with
RT   specific tyrosines on the receptor that are not required for a
RT   mitogenic response.";
RL   Mol. Cell. Biol. 16:1622-1631(1996).
RN   [14]
RP   INTERACTION WITH CRKL AND LYN.
RX   MEDLINE=21413951; PubMed=11443118; DOI=10.1074/jbc.M102924200;
RA   Arai A., Kanda E., Nosaka Y., Miyasaka N., Miura O.;
RT   "CrkL is recruited through its SH2 domain to the erythropoietin
RT   receptor and plays a role in Lyn-mediated receptor signaling.";
RL   J. Biol. Chem. 276:33282-33290(2001).
RN   [15]
RP   INTERACTION WITH APS.
RX   MEDLINE=22510236; PubMed=12444928; DOI=10.1042/BJ20020716;
RA   Wollberg P., Lennartsson J., Gottfridsson E., Yoshimura A.,
RA   Ronnstrand L.;
RT   "The adapter protein APS associates with the multifunctional docking
RT   sites Tyr-568 and Tyr-936 in c-Kit.";
RL   Biochem. J. 370:1033-1038(2003).
RN   [16]
RP   PHOSPHORYLATION, AND STAT5 ACTIVATION.
RX   PubMed=11290583; DOI=10.1182/blood.V97.8.2230;
RA   Barber D.L., Beattie B.K., Mason J.M., Nguyen M.H.-H., Yoakim M.,

RA   Neel B.G., D'Andrea A.D., Frank D.A.;
RT   "A common epitope is shared by activated signal transducer and
RT   activator of transcription-5 (STAT5) and the phosphorylated
RT   erythropoietin receptor: implications for the docking model of STAT
RT   activation.";
RL   Blood 97:2230-2237(2001).
RN   [17]
RP   C-GLYCOSYLATION, MUTAGENESIS OF SER-233 AND ALA-234, AND
RP   CARBOHYDRATE-LINKAGE SITE TRP-232.
RX   PubMed=12859190; DOI=10.1021/bi034112p;
RA   Furmanek A., Hess D., Rogniaux H., Hofsteenge J.;
RT   "The WSAWS motif is C-hexosylated in a soluble form of the
RT   erythropoietin receptor.";
RL   Biochemistry 42:8452-8458(2003).
CC   -!- FUNCTION: Receptor for erythropoietin. Mediates erythropoietin-
CC   induced erythroblast proliferation and differentiation. Upon EPO
CC   stimulation, EPOR dimerizes triggering the JAK2/STAT5 signaling
CC   cascade. In some cell types, can also activate STAT1 and STAT3.
CC   May also activate the LYN tyrosine kinase.
CC   -!- SUBUNIT: Forms homodimers on EPO stimulation. The tyrosine-
CC   phosphorylated form interacts with several SH2 domain-containing
CC   proteins including LYN, the adapter protein APS, PTPN6, PTPN11,
CC   JAK2, PI3 kinases, STAT5A/B, SOCS3 and CRKL. The N-terminal SH2
CC   domain of PTPN6 binds Tyr-453 and inhibits signaling through SH2
CC   dephosphorylation of JAK2. APS binding also inhibits the JAK-STAT
CC   signaling. Binding to PTPN11, preferentially through the N-
CC   terminal SH2 domain, promotes mitogenesis and phosphorylation of
CC   PTPN11. Binding of JAK2 (through its N-terminal) promotes cell-
CC   surface expression. Interaction with the ubiquitin ligase NOSIP
CC   mediates EPO-induced cell proliferation. Interacts with ATXN2L (By
CC   similarity).
CC   -!- SUBCELLULAR LOCATION: Type I membrane protein (isoform 1).
CC   Secreted (isoform 2).
CC   -!- ALTERNATIVE PRODUCTS:
CC   Event=Alternative splicing; Named isoforms=2;
CC   Comment=Additional isoforms seem to exist;
CC   Name=EPOR-F; Synonyms=Membrane-bound form;
CC   IsoId=P14753-1; Sequence=Displayed;
CC   Name=EPOR-S; Synonyms=Soluble form;
CC   IsoId=P14753-2; Sequence=VSP_009512, VSP_009513;
CC   -!- TISSUE SPECIFICITY: Expressed in relatively mature erythroid
CC   progenitor cells and in EPO-responsive erythroleukemia cells.
CC   -!- DOMAIN: The WSXWS motif appears to be necessary for proper protein
CC   folding and thereby efficient intracellular transport and cell-
CC   surface receptor binding.
CC   -!- DOMAIN: The box 1 motif is required for JAK interaction and/or
CC   activation.

Query Match   80.1%; Best Local Similarity 81.7%; Score 2174.5; DB 1; Length 507;
Matches 415; Conservative 24; Mismatches 68; Indels 1; Gaps 1;

```
Qy    1 MDHLGASLWPQVGSLCLLLAGAAWAPPPNLPDPKFESKAALLAARGPEELLCFTERLEDL 60
          ||:| ||:||:|| ||||||||||| | :|||||||||||||||:||:||||||:||||
Db    1 MDKLRVPLWPRVGPLCLLLAGAAWAPSPSLPDPKFESKAALLASRGSEELLCFTQRLEDL 60

Qy   61 VCFWEEAASAGVGPGNYSFSYQLEDEPWKLCRLHQAPTARGAVRFWCSLPTADTSSFVPL 120
          |||||||||:|  ::|:YSFSYQLE: :|:| | :||||:||:||||||||||||||||
Db   61 VCFWEEAASSGM-DFNYSFSYQLEGESRKSCSLHQAPTVRGSVRFWCSLPTADTSSFVPL 119

Qy  121 ELRVTAASGAPRYHRVIHINEVVLLDAPVGLVARLADESGHVVLRWLPPPETPMTSHIRY 180
          ||:||:||| ||||||:|||||||||||:||:|| | :|:GHVVLRWLPP  :PMT:HIRY
Db  120 ELQVTEASGSPRYHRIIHINEVVLLDAPAGLLARRAEEGSHVVLRWLPPGAPMTTHIRY 179

Qy  181 EVDVSAGNGAGSVQRVEILEGRTECVLSNLRGRTRYTFAVRARMAEPSFGGFWSAWSEPV 240
          |||||||| ||:.|||:|:|||||||||||||.|||||||||||||||||:|FWSAWSEP|
Db  180 EVDVSAGNRAGGTQRVEVLEGRTECVLSNLRGGTRYTFAVRARMAEPSFSGFWSAWSEPA 239

Qy  241 SLLTPSDLDPLILTLSLILVLVLLTVLALLSHRRALKQKIWPGIPSPESEFEGLFTTH 300
          ||||.||||||||||||||||:.|||||||||||||::||||:|WGIPSPESEFEGLFTTH
Db  240 SLLTASDLDPLILTLSLILVLISLLLTVLALLSHRRTLQQKIWGIPSPESEFEGLFTTH 299

Qy  301 KGNFQLWLYQNDCCLWWSPCTPFTEDPPASLEVLSERCWGTMQAVEPGTDDEGPLLEPVG 360
```

```
Db   300  |||||| || |||||||      |||||||| ||||| : |  ||||| ||||||   ||||||||||||||
          KGNFQLWLLQRDGCLWWSPGSSFPEDPPAHLEVLSEPRWAVTQAGDPGADDEGPLLEPVG  359

Qy   361  SEHAQDTYLVLDKWLLPRNPPSEDLPGPGGSVDIVAMDEGSEASSCSSALASKPSPEGAS  420
          |||||||||||||||||||  |:| ||||||||| |   |||| :|| ||||
Db   360  SEHAQDTYLVLDKWLLPRTPCSENLSGPGGSVDPVTMDEASETSSCPSDLASKPRPEGTS  419

Qy   421  AASFEYTILDPSSQLLRPWTLCPELPPTPHLKYLYLVVSDSGISTDYSSSGDSQAQGGL  480
          :|||| |||||||||||||   ||||:||||||||||||||||||||||  ||  |
Db   420  PSSFEYTILDPSSQLLCPRALPPELPPTPHLKYLYLVVSDSGISTDYSSSGGSQGVHGDS  479

Qy   481  SDGPYSNPYENSLIPAAEPLPPSYVACS  508
          |||||||||||| :|||:  : :||  |
Db   480  SDGPYSHPYENSLVPDSEPLHPGYVACS  507
```

RESULT 4
EPOR_RAT

```
ID   EPOR_RAT     STANDARD;      PRT;   507 AA.
AC   Q07303;
DT   01-FEB-1995 (Rel. 31, Created)
DT   01-FEB-1995 (Rel. 31, Last sequence update)
DT   25-OCT-2004 (Rel. 45, Last annotation update)
DE   Erythropoietin receptor precursor (EPO-R).
GN   Name=Epor;
OS   Rattus norvegicus (Rat).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX   NCBI_TaxID=10116;
RN   [1]
RP   SEQUENCE FROM N.A. (ISOFORM EPOR-F).
RC   STRAIN=Long Evans;
RX   PubMed=9029168; DOI=10.1016/S0304-3835(96)04544-2;
RA   Fujita M., Takahashi R., Kitada K., Watanabe R., Kitazawa S.,
RA   Ashoori F., Liang P., Saya H., Serikawa T., Maeda S.;
RT   "Alternative splicing of the erythropoietin receptor gene correlates
RT   with erythroid differentiation in rat hematopoietic and leukemic
RT   cells.";
RL   Cancer Lett. 112:47-55(1997).
RN   [2]
RP   SEQUENCE FROM N.A. (ISOFORMS EPOR-F AND EPOR-S).
RC   STRAIN=Long Evans;
RC   TISSUE=Pheochromocytoma;
RX   MEDLINE=9326574; PubMed=7684373;
RA   Masuda S., Nagao M., Takahata K., Konishi Y., Gallyas F., Tabira T.,
RA   Sasaki R.;
RT   "Functional erythropoietin receptor of the cells with neural
RT   characteristics. Comparison with receptor properties of erythroid
RT   cells.";
RL   J. Biol. Chem. 268:11208-11216(1993).
CC   -!- FUNCTION: Receptor for erythropoietin. Mediates erythropoietin-
CC       induced erythroblast proliferation and differentiation. Upon EPO
CC       stimulation, EPOR dimerizes triggering the JAK2/STATS signaling
CC       cascade. In some cell types, can also activate STAT1 and STAT3.
CC       May also activate LYN tyrosine kinase (By similarity).
CC   -!- SUBUNIT: Forms homodimers on EPO stimulation. The tyrosine-
CC       phosphorylated form interacts with several SH2 domain-containing
CC       proteins including LYN, the adapter protein APS, PTPN6, PTPN11,
CC       JAK2, PI3 kinases, STAT5A/B, SOCS3 and CRKL. The N-terminal SH2
CC       domain of PTPN6 binds Tyr-453 and inhibits signaling through
CC       dephosphorylation of JAK2. APS binding also inhibits the JAK-STAT
CC       signaling. Binding to PTPN11, preferentially through the N-
CC       terminal SH2 domain, promotes mitogenesis and phosphorylation of
CC       PTPN11. Binding of JAK2 (through its N-terminal) promotes cell-
CC       surface expression. Interaction with the ubiquitin ligase NOSIP
CC       mediates EPO-induced cell proliferation. Interacts with ATXN2L (By
CC       similarity).
CC   -!- SUBCELLULAR LOCATION: Type I membrane protein.
CC   -!- ALTERNATIVE PRODUCTS:
CC       Event=Alternative splicing; Named isoforms=2;
CC       Name=EPOR-F; Synonyms=Full-length form;
CC         IsoId=Q07303-1; Sequence=Displayed;
CC       Name=EPOR-S; Synonyms=Soluble form;
CC         IsoId=Q07303-2; Sequence=VSP_009514, VSP_009515;
CC   -!- TISSUE SPECIFICITY: Both isoforms expressed in bone marrow, spleen
CC       and erythroleukemia cell lines.
CC   -!- DOMAIN: The WSXWS motif appears to be necessary for proper protein
CC       folding and thereby efficient intracellular transport and cell-
CC       surface receptor binding.
CC   -!- DOMAIN: The box 1 motif is required for JAK interaction and/or
CC       activation.
CC   -!- DOMAIN: Contains 1 copy of a cytoplasmic motif that is referred to
CC       as the immunoreceptor tyrosine-based inhibitor motif (ITIM). This
CC       motif is involved in modulation of cellular responses. The
CC       phosphorylated ITIM motif can bind the SH2 domain of several SH2-
CC       containing phosphatases.
CC   -!- PTM: On EPO stimulation, phosphorylated on C-terminal tyrosine
CC       residues by JAK2. The phosphotyrosine motifs are also recruitment
CC       sites for several SH2-containing proteins and adapter proteins
CC       which mediate cell proliferation. Phosphorylation on Tyr-453 is
CC       required for PTPN6 interaction, Tyr-425 for PTPN11. Tyr-425 is
CC       also required for SOCS3 binding, but Tyr-453/Tyr-455 motif is the
CC       preferred binding site (By similarity).
CC   -!- PTM: Ubiquitinated by NOSIP; appears to be either multi-
CC       monoubiquitinated or polyubiquitinated. Ubiquitination mediates
CC       proliferation and survival of EPO-dependent cells (By similarity).
CC   -!- SIMILARITY: Belongs to the type I cytokine family of receptors.
CC       Subfamily 1.
CC   -!- SIMILARITY: Contains 1 fibronectin type III domain.
CC   -------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   -------------------------------------------------------------------------
DR   EMBL; D13566; BAA02761.1; -.
DR   PIR; A46713; A46713.
DR   HSSP; P19235; 1EBA.
DR   RGD; 2560; Epor.
DR   InterPro; IPR002996; Cytkn_recept_B/G.
DR   InterPro; IPR009167; EPO_receptor.
DR   InterPro; IPR003961; FN_III.
DR   InterPro; IPR008957; FN_III-like.
DR   InterPro; IPR003528; Hemptrecept1_F1.
DR   Pfam; PF00041; fn3; 1.
DR   PIRSF; PIRSF001959; EPO_receptor; 1.
DR   PROSITE; PS50853; FN3; 1.
DR   PROSITE; PS01352; HEMATOPO_REC_L_F1; 1.
KW   Alternative splicing; Glycoprotein; Phosphorylation; Receptor; Signal;
KW   Transmembrane; Ubl conjugation.
FT   SIGNAL        1     24                                          By similarity.
FT   CHAIN        25    507       Erythropoietin receptor.
FT   DOMAIN       25    249       Extracellular (Potential).
FT   TRANSMEM    250    272       Potential.
FT   DOMAIN      273    507       Cytoplasmic (Potential).
FT   DOMAIN      143    239       Fibronectin type-III.
FT   SITE        116    116       Required for ligand binding (By
FT                                similarity).
FT   SITE        232    236       WSXWS motif.
FT   SITE        281    289       Box 1 motif.
FT   SITE        367    367       APS binding, and STAT5 binding and
FT                                activation.
FT   SITE        425    425       Required for STAT5/PTPN11/SOCS3 binding.
FT   SITE        451    456       ITIM motif.
FT   SITE        453    453       PTPN6 binding.
FT   DISULFID     52     62       By similarity.
FT   DISULFID     90    106       By similarity.
FT   MOD_RES     367    367       Phosphotyrosine (by JAK2) (By
FT                                similarity).
FT   MOD_RES     425    425       Phosphotyrosine (by JAK2) (By
FT                                similarity).
FT   MOD_RES     453    453       Phosphotyrosine (by JAK2) (By
FT                                similarity).
FT   MOD_RES     455    455       Phosphotyrosine (by JAK2) (By
```

```
FT
FT   MOD_RES     467    467                                    similarity).
FT                                                             Phosphotyrosine (by JAK2) (By
FT   MOD_RES     484    484                                    similarity).
FT                                                             Phosphotyrosine (by JAK2) (By
FT   MOD_RES     484    484                                    similarity).
FT                                                             Phosphotyrosine (by JAK2) (By
FT   MOD_RES     488    488                                    similarity).
FT                                                             Phosphotyrosine (by JAK2) (By
FT   MOD_RES     503    503                                    similarity).
FT                                                             Phosphotyrosine (by JAK2) (By
FT   CARBOHYD     75     75                                    N-linked (GlcNAc...) (Potential).
FT   VARSPLIC    246    265                                    DLDPLILTLSLILVLISLLL -> GEATVPRGGGGAGPNTS
FT                                                             KPP (in isoform EPOR-S).
FT                                                             /FTId=VSP_009514.
FT   VARSPLIC    266    507                                    Missing (in isoform EPOR-S).
FT                                                             /FTId=VSP_009515.
SQ   SEQUENCE    507 AA;  55499 MW;  AC79AF22D06A7312 CRC64;

Query Match          79.3%;  Score 2152.5;  DB 1;  Length 507;
Best Local Similarity 80.5%;  Pred. No. 8.2e-142;
Matches 409;  Conservative 31;  Mismatches 67;  Indels 1;  Gaps 1;

Qy     1 MDHLGASLWPQVGSLCLLLAGAWAPPNLPDDKPFESKAALLAARGPELLCFTERLEDL  60
Db     1 MDQLRVARWPRVSPLCLLLAGAWASSPSLPDDPKFESKAALLASRGSELLCFTQRLEDL  60

Qy    61 VCFWEEAASAGVGPGNYSFSYQLEDEPWKLCRLHQAPTARGAVRFWCSLPTADTSSFVPL 120
Db    61 VCFWEEAANSGMG-FNYSFSYQLEGESRKSCRLHQAPTVRGSMRFWCSLPTADTSSFVPL 119

Qy   121 ELRVTAASGAPRYRVHINEVLLDAPVGLVARLADESGHVVLRWLPPPETPMTSHIRY    180
Db   120 ELQVTEASGSSPRYRHIIHINEVLLDAPAGLLARRAEBGSHVVLRWLRWLPPGAPMTTHIRY 179

Qy   181 EVDVSAGNGAGSVQRVEILEGRTECVLSNLRGRTRYTFAVRARMAEPSFGGFWSAWSEPV 240
Db   180 EVDVSAGNRAGGTQRVEVLEGRTECVLSNLRGGTRYTFAVRARMAEPSFGGFWSAWSEPA 239

Qy   241 SLLTPSDLDPLIITLSLIIVILVILLTVLALLSHRRALKQKIWPGIPSPESEFEGLFTTH 300
Db   240 SLLTASDLDPLIITLSLIIVLILVLISLLLTVLALLSHRRALRQKIWPGIPSPENEFEGLFTTH 299

Qy   301 KGNFQLWLYQNDGCLWWSPCTPFTEDPPASLEVLSERCWGTMQAVEPGTDDEGPLLEPVG 360
Db   300 KGNFQLWLLQRDGCLWWSPSSFPFEDPPAHLEVLSERRWGVTQAGDAGAEDKGPLLEPVG 359

Qy   361 SEHAQDTYLVLDKWLLPRNPPSEDLPGPGGSVDIVAMDEGSEASSCSSALASKPSPEGAS 420
Db   360 SERAQDTYLVLDEWLLPRCPCSENLSGPGDSVDPATMDEGSETSSCPSDLASKPRPEGTS 419

Qy   421 AASFEYTILDPSSQLLRPWTLCPELPPTPPHLKYLYLVVSDSGISTDYSSSGDSQGAQGGL 480
Db   420 PSSFEYTILDPSSKLLCPRALPPELPPTPPHLKYLYLVVSDSGISTDYSSSGGSQGVHGDS 479

Qy   481 SDGPYSNPYENSLIPAAEPLPPSYVACS 508
Db   480 SDGPYSHPYENSLVPDTEPLRPSYVACS 507

RESULT 5
Q95N13
ID   Q95N13;     PRELIMINARY;     PRT;     387 AA.
AC   Q95N13;
DT   01-DEC-2001 (TrEMBLrel. 19, Created)
DT   01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT   01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE   Erythropoietin receptor (Fragment).
OS   Ovis aries (Sheep).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC   Caprinae; Ovis.
OX   NCBI_TaxID=9940;
RN   [1]
RP   SEQUENCE FROM N.A.
```

```
RA   David B., Lim G.B., Moritz K.M., Koukoulas I., Wintour M.;
RL   Submitted (APR-2001) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AY029232; AAK38737.1; -.
DR   HSSP; P19235; 1EBA.
DR   GO; GO:0004872; F:receptor activity; IEA.
DR   InterPro; IPR003961; FN_III.
DR   InterPro; IPR008957; FN_III-like.
DR   InterPro; IPR003528; Hemptreceptl_F1.
DR   Pfam; PF00041; fn3; 1.
DR   SMART; SM00060; FN3; 1.
DR   PROSITE; PS50853; FN3; 1.
DR   PROSITE; PS01352; HEMATOPO_REC_L_F1; UNKNOWN_1.
KW   Receptor.
FT   NON_TER       1      1
FT   NON_TER     387    387
SQ   SEQUENCE    387 AA;  42039 MW;  0D1E6173C432EBC6 CRC64;

Query Match          61.5%;  Score 1670.5;  DB 2;  Length 387;
Best Local Similarity 82.9%;  Pred. No. 2.4e-108;
Matches 319;  Conservative 14;  Mismatches 51;  Indels 51;  Gaps 1;

Qy    68 ASAGVGPGNYSFSYQLEDEPWKLCRLHQAPTARGAVRFWCSLPTADTSSFVPLELRVTAA 127
Db     3 ATAGVGPDNYSFSYQLEGEPWKPCRLHQTPTARGLVRFWCSLPTADTSSFVPLELHVTAA  62

Qy   128 -SGAPRYRVHINEVLLDAPVGLVARLADESGHVVLRWLPPPETPMTSHIRYEVDVSA 186
Db    63 SSGASRYRRTIHVNEVLLDPPARLVARRADEGGHVVLRWLPPPGAPMASLIRYEVNISA 122

Qy   187 GNGAGSVQRVEILEGRTECVLSNLRGRTRYTFAVRARMAEPSFGGFWSAWSEPVSLLTPS 246
Db   123 ENAAGGAQRVEILDGRTECLLSNLRGGTRYTFMVRARMAEPSFGGFWSAWSEPASLLTAS 182

Qy   247 DLDPLILTLSLIIVILVLLTVLALLSHRRALKQKIWPGIPSPESEFEGLFTTHKGNFQL 306
Db   183 DLDPLILTLSLIVLVLILLLAVLALLSHRRTLKQKIWPGIPSPESEFEGLFTTHKGNFQL 242

Qy   307 WLYQNDGCLWWSPCTPFTEDPPASLEVLSERCWGTMQAVEPGTDDEGPLLEPVGSEHAQD 366
Db   243 WLYQTDGCLWWSPSTPFPEDPPAPLEVLSECCWGVTQAVEPGADDGGSLLEPVGSEHAQD 302

Qy   367 TYLVLDKWLLPRNPPSEDLPGPGGSVDIVAMDEGSEASSCSSALAKPSPEGASAASFEY 426
Db   303 SYLMLDKWLLPRSPPSEDLLQPCGDMDIVTTDEASEASSCTSALALKPGEGASAASFEY 362

Qy   427 TILDPSSQLLRPWTLCPELPPTPPH 451
Db   363 TILDPSSQLLRPRALPPELLPTPPH 387

RESULT 6
Q95N14
ID   Q95N14;     PRELIMINARY;     PRT;     418 AA.
AC   Q95N14;
DT   01-DEC-2001 (TrEMBLrel. 19, Created)
DT   01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT   01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DB   Erythropoietin receptor (Fragment).
OS   Ovis aries (Sheep).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC   Caprinae; Ovis.
OX   NCBI_TaxID=9940;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   David B., Lim G.B., Moritz K.M., Koukoulas I., Wintour M.;
RL   Submitted (APR-2001) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AY029231; AAK38170.1; -.
DR   HSSP; P19235; 1EBA.
DR   GO; GO:0004872; F:receptor activity; IEA.
DR   InterPro; IPR003961; FN_III.
DR   InterPro; IPR008957; FN_III-like.
DR   InterPro; IPR003528; Hemptreceptl_F1.
```

```
DR   Pfam; PF00041; fn3; 1.
DR   SMART; SM00060; FN3; 1.
DR   PROSITE; PS50853; FN3; 1.
DR   PROSITE; PS01352; HEMATOPO_REC_L_F1; UNKNOWN_1.
KW   Receptor.
FT   NON_TER       1
FT   NON_TER     418
SQ   SEQUENCE   418 AA;  45282 MW;  BB742EBEA03450C CRC64;
```

```
Query Match             60.6%;  Score 1645;  DB 2;  Length 418;
Best Local Similarity   76.7%;  Pred. No. 1.6e-106;
Matches  319;  Conservative  14;  Mismatches  51;  Indels  32;  Gaps  2;

Qy   68  ASAGVGPGNYSFSYQLEDEPWKLCRLHQAPTARGAVRFWCSLPTADTSSFVPLELRVTAA 127
             |||||||:||||||||||:||||| |||:|||||:|||||||||||||||||| ||||
Db    3  ATAGVGPDNYSFSYQLEGEPWKPCRLHQTPTARGLVRFWCSLPTADTSSFVPLELHVTAA  62

Qy  128  -SGAPRYHRVIHINEVVLLDAPVGLVARLADESGHVVLRWLPPPETPMTSHIRYEVDVSA 186
             |||:|||||:||:|||||| ||| ||||:||:|||||||||| ||| :||||||::|
Db   63  SSGASRYRRTIHVNEVVLLDPPARLVARRADEGGHVVLRWLPPPGAPMASLIRYEVNISA 122

Qy  187  GNGAGSVQRVEILEGRTECVLSNLRGRTRYTFAVRARMAEPSFGGFWSAWSEPVSLLTPS 246
             :|||| ||||||:|||||:||||||| ||||||||||||||||||||||||||:||| |
Db  123  ENAAGGAQRVEILDGRTECLLSNLRGGTRYTFMVRARMAEPSFGGFWSAWSEPASLLTAS 182

Qy  247  DLDPLILTLSLIIVILVILLTVLALLSHRRALKQKIWPGIPSPSEFEGLFTTHKGNF-- 304
             |||||||||||:|||| || ||||||||||:||||||||||||||||||||||||||
Db  183  DLDPLILTLSLVLVILLLLAVLALLSHRRTLKQKIWPGIPSPSEFEGLFTTHKGNFPQV 242

Qy  305  -----------QLWLYQNDGCLWWSPCTPFTEDPPASLEVLS 335
                      ||||| :||||||||| |||:|||||:|||||
Db  243  GGLVVPSRPGASLLLWPNSKSLSKLJVLFPQLWLYQTDGCLWWSPTPFPEDPPAPLEVLS 302

Qy  336  ERCWGTMQAVEPGTDDEGPLLEPVGSEHAQDTYIVLDKWLLPRNPPSEDLPGPGGSVDIV 395
             | |||:|||||||:|::|||||||||||||:|:::||||||||:|||:|||  |: ||||
Db  303  ECCWGVTQAVEPGADDGGSLLEPVGSEHAQDSYLMLDKWLLPRSPPSSEDLLQPCGDMDIV 362

Qy  396  AMDEGSEASSCSSALASKPSPEGASASPFEYTILDPSSQLLRPWTLCPELPPTPPH 451
             ::|| ||||||:|||| || |||||| |||||||||||||||||  || |||||||
Db  363  TTDEASEASSCTSALALKPGPEGASAASPFEYTILDPSSQLLRPRALPPELLPTPPH 418
```

```
RESULT 7
O35545
ID   O35545      PRELIMINARY;     PRT;   316 AA.
AC   O35545;
DT   01-JAN-1998 (TrEMBLrel. 05, Created)
DT   01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Intron 5-inserted form of erythropoietin receptor precursor.
OS   Rattus norvegicus (Rat).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX   NCBI_TaxID=10116;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=98296111; PubMed=9630610; DOI=10.1016/S0167-4889(98)00037-8;
RA   Yamaji R., Murakami C., Takenoshita M., Tsuyama S., Inui H.,
RA   Miyatake K., Nakano Y.;
RT   "The intron 5-inserted form of rat erythropoietin receptor is
RT   expressed as a membrane-bound form.";
RL   Biochim. Biophys. Acta 1403:169-178(1998).
EMBL; D83509; BAA22373.1; -.
HSSP; P19235; 1EBA.
NCBI_TaxID=10116;
```

```
DR   GO; GO:0016020; C:membrane; IEA.
DR   GO; GO:0004896; F:hematopoietin/interferon-class (D200-domain. . .; IEA
DR   GO; GO:0004872; F:receptor activity; IEA.
DR   InterPro; IPR002996; Cytkn_recept_B/G.
DR   InterPro; IPR003961; FN_III.
DR   InterPro; IPR008957; FN_III-like.
DR   InterPro; IPR003528; Hemptreceptl_F1.
DR   Pfam; PF00041; fn3; 1.
DR   SMART; SM00060; FN3; 1.
```

```
DR   PROSITE; PS50853; FN3; 1.
DR   PROSITE; PS01352; HEMATOPO_REC_L_F1; 1.
KW   Receptor; Signal.
FT   SIGNAL        1     24       Potential.
FT   CHAIN        25    316       Potential.
SQ   SEQUENCE   316 AA;  34220 MW;  05C44BF8516C180B CRC64;
```

```
Query Match             39.0%;  Score 1058;  DB 2;  Length 316;
Best Local Similarity   68.6%;  Pred. No. 9.9e-66;
Matches  212;  Conservative  23;  Mismatches  60;  Indels  14;  Gaps  3;

Qy    1  MDHLGASLWPQVGSLCLLLAGAAWAPPNLPDPKFESKAALLAARGPEELLCFTERLEDL  60
             ||::||||  ||| |||||||||||  ::||: |||||||||||| ||:||||| ||||||
Db    1  MDQLRVARWPRVSPLCLLLAGAAWASSPSLPDDPKFESKAALLASRGSEELLCFTQRLEDL  60

Qy   61  VCFWEEAASAGVGPGNYSFSYQLEDEPWKLCRLHQAPTARGAVRFWCSLPTADTSSFVPL 120
             |||||||::  |:||||   |||||||||:   |:|:|||||||:||:||||||||||||||
Db   61  VCFWEEAANSGMG-FNYSFSYQLEGESRKSCRLHQAPTVRGSMRFWCSLPTADTSSFVPL 119

Qy  121  ELRVTAASGAPRYHRVIHINEVVLLDAPVGLVARLADESGHVVLRWLPPPETPMTSHIRY 180
             ||:|| ||||:|||:|||||:||||||||:|||: :|: ||||||||||| ||:||:|||
Db  120  ELQVTEASGSPRYRHIIHINEVVLLDAPAGLLARRAEEGSHVVLRWLPPPGAPMTHIRY 179

Qy  181  EVDVSAGNGAGSVQRVEILEGRTECVLSNLRGRTRYTFAVRARMAEPSFGGFWSAWSEPV 240
             ||||||||:||  ||||:||||||||||||||| ||||||||||||||||||||||||||:
Db  180  EVDVSAGNRAGGTQRVEVLEGRTECVLSNLRGGTRYTFAVRARMAEPSFGGFWSAWSEPA 239

Qy  241  SLLTPS----------DLDPLILTLSLIIVILVILLTVLALLSHRR---ALKQKIWPGIP 287
             ||||                                     |:|
Db  240  SLLTASGEATVPRGGGAGPNTRQTSLTFLPPRLGPSHLDAVSHSRPHLTVADCAGPAVP 299

Qy  288  SPESEFEGL 296
             ||||||
Db  300  PPGSAAEDL 308
```

```
RESULT 8
Q27950
ID   Q27950      PRELIMINARY;     PRT;   229 AA.
AC   Q27950;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE   Erythropoietin receptor (Fragment).
GN   Name=EpoR;
OS   Bos indicus (Zebu).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC   Bovinae; Bos.
OX   NCBI_TaxID=9915;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Bone marrow;
RA   Suliman H.B., Feldman B. F, Majiwa P.A.O., Logan-Henfrey L.L.;
RL   Submitted (JUN-1996) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; U61398; AAB03870.1; -.
DR   HSSP; P19235; 1EBA.
DR   GO; GO:0004872; F:receptor activity; IEA.
DR   InterPro; IPR002996; Cytkn_recept_B/G.
DR   InterPro; IPR003961; FN_III.
DR   InterPro; IPR008957; FN_III-like.
DR   InterPro; IPR003528; Hemptreceptl_F1.
DR   Pfam; PF00041; fn3; 1.
DR   SMART; SM00060; FN3; 1.
DR   PROSITE; PS50853; FN3; 1.
DR   PROSITE; PS01352; HEMATOPO_REC_L_F1; UNKNOWN_1.
KW   Receptor.
FT   NON_TER       1
FT   NON_TER     229
SQ   SEQUENCE   229 AA;  25196 MW;  F6E01C4AB07893E8 CRC64;
```

```
Query Match             37.4%;  Score 1016;  DB 2;  Length 229;
Best Local Similarity   84.7%;  Pred. No. 5.6e-63;
```

```
Matches 194; Conservative 9; Mismatches 26; Indels 0; Gaps 0;

Qy  59 DLVCFWEEAASAGVGPGNYSFSYQLEDEPWKLCRLHQAPTARGAVRFWCSLPTADTSSFV 118
        ||||||||||| |||||| ||||||||||||| |||||||||| ||||||||||||||||
Db   1 DLVCFWEEAAATAGVGPDNYSFSYQLEGEPWKPCRLHQAPTARGLVRFWCSLPTADTSSFV 60

Qy 119 PLELRVTAASGAPRYHRVIHINEVVLLDAPVGLVARLADESGHVVLRWLPPETPMTSHI 178
        |||| |||||| ||| |:|| ||||| ||| |||| ||| |||||||||
Db  61 PLELHVTAASGASRYRRTIHVNEVVLLDPPARLVARRADEGHVVLRWLPPPGAPMASLI 120

Qy 179 RYEVDVSAGNGAGSVQRVEILEGRTECVLSNLRGRTRYTFAVRARMAEPSFGGFWSAWSE 238
        ||||:|| ||||||:||||||||||||||||||| ||||| ||||||||||||||||||
Db 121 RYEVNISAENAAGSAQRVEILDGRTECLLSNLRGGTRYTFMVRARMAEPSFGGFWSAWSE 180

Qy 239 PVSLLTPSDLDPLILTLSLILVVILVILLTVLALLSHRRALKQKIWPGIP 287
        |||||| ||||||||||||:|:|:|| ||||:|| |||||||||
Db 181 PASLLTASDLDPLILTLSLVLVLILLLAVLALLSHRRTLKQKIWPGIP 229


RESULT 9
Q28206
ID   Q28206;     PRELIMINARY;     PRT;   229 AA.
AC   Q28206;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE   Erythropoietin receptor (Fragment).
GN   Name=EpoR;
OS   Bos taurus (Bovine).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC   Bovinae; Bos.
OX   NCBI_TaxID=9913;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Bone marrow;
RA   Suliman H.B., Feldman B. F., Majiwa P.A.O., Logan-Henfrey L.L.;
RL   Submitted (JUN-1996) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; U61399; AAB03871.1; -.
DR   HSSP; P19235; 1EBA.
DR   GO; GO:0004872; F:receptor activity; IEA.
DR   InterPro; IPR002996; Cytkn_recept_B/G.
DR   InterPro; IPR003961; FN III.
DR   InterPro; IPR0008957; FN III-like.
DR   InterPro; IPR003528; Hemptreceptl_F1.
DR   Pfam; PF00041; fn3; 1.
DR   SMART; SM00060; FN3; 1.
DR   PROSITE; PS50853; FN3; 1.
DR   PROSITE; PS01352; HEMATOPO_REC_L_F1; UNKNOWN_1.
KW   Receptor.
FT   NON_TER        1      1
FT   NON_TER      229    229
SQ   SEQUENCE   229 AA;  25170 MW;  EDFAA6F11OD992E8 CRC64;

Query Match             37.2%;  Score 1011;  DB 2;  Length 229;
Best Local Similarity   84.3%;  Pred. No. 1.3e-62;
Matches 193;  Conservative 10;  Mismatches 26;  Indels 0;  Gaps 0;

Qy  59 DLVCFWEEAASAGVGPGNYSFSYQLEDEPWKLCRLHQAPTARGAVRFWCSLPTADTSSFV 118
        ||||||||||| |||||| ||||||||||||| |||||||||| ||||||||||||||||
Db   1 DLVCFWEEAAATAGVGPDNYSFSYQLEGEPWKPCRLHQAPTARGLVRFWCSLPTADTSSFV 60

Qy 119 PLELRVTAASGAPRYHRVIHINEVVLLDAPVGLVARLADESGHVVLRWLPPETPMTSHI 178
        |||| |||||| ||| |:|| ||||| ||| |||| ||| |||||||||
Db  61 PLELHVTAASGASRYRRTIHVNEVVLLDPPARLVARRADEGHVVLRWLPPPGAPMASLI 120

Qy 179 RYEVDVSAGNGAGSVQRVEILEGRTECVLSNLRGRTRYTFAVRARMAEPSFGGFWSAWSE 238
        ||||:|| ||||||:||||||||||||||||||| ||||| ||||||||||||||||||
Db 121 RYEVNISAENAAGSAQRVEILDGRTECLLSNLRGGTRYTFMVRARMAEPSFGGFWSAWSE 180

Qy 239 PVSLLTPSDLDPLILTLSLILVVILVILLTVLALLSHRRALKQKIWPGIP 287
        |||||| ||||||||||||:|:|:|| ||||:|| |||||||||
Db 181 PASLLTASDLDPLILTLSLVLVLILLLAVLALLSHRRTLKQKIWPGIP 229
```

```
RESULT 10
Q6UAP7
ID   Q6UAP7      PRELIMINARY;     PRT;   529 AA.
AC   Q6UAP7;
DT   05-JUL-2004 (TrEMBLrel. 27, Created)
DT   05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT   05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE   Class I helical cytokine receptor number 9.
GN   Name=CRFA9;
OS   Tetraodon nigroviridis (Green puffer).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC   Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes;
OC   Tetraodontoidea; Tetraodontidae; Tetraodon.
OX   NCBI_TaxID=99883;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Jaillon O., Aury J.-M., Brunet F., Petit J.-L., Stange-Thomann N.,
RA   Mauceli E., Bouneau L., Fischer C., Ozouf-Costaz C., Bernot A.,
RA   Nicaud S., Jaffe D., Fisher S., Lutfalla G., Dossat C., Segurens B.,
RA   Dasilva C., Salanoubat M., Levy M., Boudet N., Castellano S.,
RA   Anthouard V., Jubin C., Castelli V., Katinka M., Vacherie B.,
RA   Biémont C., Skalli Z., Cattolico L., Poulain J., Berardinis Vd.,
RA   Cruaud C., Duprat S., Brottier P., Coutanceau J.-P., Gouzy J.,
RA   Parra G., Lardier G., Chapple C., McKernan K.J., McEwan P., Bosak S.,
RA   Kellis M., Volff J.-N., Guigs R., Zody M.C., Mesirov J.,
RA   Lindblad-Toh K., Birren B., Nusbaum C., Kahn D., Robinson-Rechavi M.,
RA   Laudet V., Schachter V., Quétier F., Saurin W., Scarpelli C.,
RA   Wincker P., Lander E.S., Weissenbach J., Crollius H.R.;
RT   "Analysis of the Tetraodon nigroviridis genome reveals the
RT   protokaryotype of bony vertebrates and its duplication in teleost
RT   fish.";
RL   Nature 0:0-0(2004).
DR   EMBL; AY374481; AAR25672.1; -.
DR   GO; GO:0004872; F:receptor activity; IEA.
DR   InterPro; IPR002996; Cytkn_recept_B/G.
DR   InterPro; IPR003961; FN III.
DR   InterPro; IPR0008957; FN_III-like.
DR   InterPro; IPR003528; Hemptreceptl_F1.
DR   Pfam; PF00041; fn3; 1.
DR   SMART; SM00060; FN3; 1.
DR   PROSITE; PS50853; FN3; 1.
DR   PROSITE; PS01352; HEMATOPO_REC_L_F1; UNKNOWN_1.
KW   Receptor.
SQ   SEQUENCE   529 AA;  59609 MW;  C4A04BFC94E9C3B6 CRC64;

Query Match             19.5%;  Score 529.5;  DB 2;  Length 529;
Best Local Similarity   27.8%;  Pred. No. 1.4e-28;
Matches 152;  Conservative 84;  Mismatches 199;  Indels 111;  Gaps 16;

Qy  15 LCLLLAGAWAPPNLPDPK-FESKAALLAARGEPELLCFTERLEDLVCFWEEAASAGVG 73
        |:  :|        | |   ::| ::|          ||:  ::||:||||||::
Db  10 LLLTIIFGATRPPANVQGAGQFRKKVSIMLKEDPKNPKCFAEGRKDFICFWEEDEERAGS 69

Qy  74 PGNYSFSYQLEDEPWKLCRLHQAPTARGAVRFWCSLPTADTSSFVPLELRVTAASGAPRY 133
        |  |:|:|  :: :  || : :||:  |:|:   | : |  ||:
Db  70 VDQYTFTYAYQNENSSRCPLKSISAADSKRLFICHLNR--IKMFVQMDIQV-HREGMLIH 126

Qy 134 HRVIHINEVVLLDAPVGLVARLADESGHVVLRWLPPETPMTSHIRYEVDVSAGNGAGSV 193
         | :: :||:|||.|..:| :. :  :..:|:||| |   |  :|||.||.  : :|
Db 127 NRSLLVEMVFLLDPPANVTVNTRKQGQLNVTWVPPLKYMDSMMYEVSYSAMD--SHV 184

Qy 194 QRVEILEGRTECVLSNLRGRTRYTFAVRARMAEPSFGGFWSAWSEPVSLLT-PSDLDPLI 252
         :||::  : |::::  | ::::|::|.| : : | :|:||||| ::: ||  ::|| |:|
Db 185 MQVEMVQASSELIIRGLQPGTKYEVQVRVKLDGISYSGYWSAWSDSVIETLPAELDLLI 244

Qy 253 LTLSLILVVILVLLTVLALLSHRRALKQKIWPGIPSPSEPEFEGLFTTHKGNFQLWLYQND 312
         :: |:::| :|:: : ::.:| || ::::||| ||.|.|  ||:   ::|:||:|| ::
Db 245 VSLALVILFVLIGLFLTTVMSNRRYLVKKIWPNIPTPDSKFHGLFSVVGGEFQSWLEQTS 304

Qy 313 GCLWWSPCTPFTEDPPASLEVLSERCWGTMQAVEPGTDDEGPLLEP----------- 358
```

```
Db   305  G-LWVTPVFFNTEELASSLEVLSELC---------PCPSPSAPPLPEDLKLPAAPNCVTD  355
          | ||.:|    ||:  :|||||||              |||
Qy   359  ------------VGSEHAQDTYLVLDKWLLP------RNP------------------  380
          |   |.:|      |:|
Db   356  GLVERRKVEAHSAVSEGWEVTADNQMPTDSWRGPQPNGVPCCSRSPLLESQDAYVTLSTNN  415
Qy   381  ------PSEDLPGPGGSVDIVAMDEGSEASSCSSALASKPSPEG---ASAASFEYT  427
          |:|    ||||     :|:||
Db   416  QREEFPLNHIPEETLP----IEKLFTSRKQICESHSDLGSMQQSPASSHLSSQSSFE--  468
          |  ||:         ::|  |||:|
Qy   428  ILDPSSQLLRPWTLCPELPPTPHLKYLYLVVSDGSGISTDYSSGDSQGAQGGLSDGPYSN  487
          :   |||::|          ||:||
Db   469  ------------CPNY--TWMSKGYVHMAAVDSGVSMDYS------PMQRVAD-VHNN  505
Qy   488  PYENSL  493
          :|:|
Db   506  EYKNGI  511


RESULT 11
IL9R_HUMAN
ID   IL9R_HUMAN     STANDARD;     PRT;    522 AA.
AC   Q01113; Q14634; Q8WWU1; Q96TF0;
DT   01-APR-1993 (Rel. 25, Created)
DT   28-FEB-2003 (Rel. 41, Last sequence update)
DT   05-JUL-2004 (Rel. 44, Last annotation update)
DE   Interleukin-9 receptor precursor (IL-9R).
GN   Name=IL9RX; Synonyms=IL9R;
GN   and
GN   Name=IL9RY; Synonyms=IL9R;
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=92302307; PubMed=1376929;
RA   Renauld J.C., Druez C., Kermouni A., Houssiau F., Uyttenhove C.,
RA   van Roost E., van Snick J.;
RT   "Expression cloning of the murine and human interleukin 9 receptor
RT   cDNAs.";
RL   Proc. Natl. Acad. Sci. U.S.A. 89:5690-5694(1992).
RN   [2]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=94250901; PubMed=8193355;
RA   Chang M.S., Engel G., Benedict C., Basu R., McNinch J.;
RT   "Isolation and characterization of the human interleukin-9 receptor
RT   gene.";
RL   Blood 83:3199-3205(1994).
RN   [3]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Melanoma;
RX   MEDLINE=96115587; PubMed=8666384;
RA   Kermouni A., van Roost E., Arden K.C., Vermeesch J.R., Weiss S.,
RA   Godelaine D., Flint J., Lurquin C., Szikora J.P., Higgs D.R.,
RA   Marynen P., Renauld J.C.;
RT   "The IL-9 receptor gene (IL9R): genomic structure, chromosomal
RT   localization in the pseudoautosomal region of the long arm of the sex
RT   chromosomes, and identification of IL9R pseudogenes at 9qter, 10pter,
RT   16pter, and 18pter.";
RL   Genomics 29:371-382(1995).
RN   [4]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=2012249; PubMed=10655549; DOI=10.1093/hmg/9.3.395;
RA   Ciccodicola A., D'Esposito M., Esposito T., Gianfrancesco F.,
RA   Migliaccio C., Miano M.G., Matarazzo M.R., Vacca M., Franze A.,
RA   Cuccurese M., Cocchia M., Curci A., Terracciano A., Torino A.,
RA   Cocchia S., Mercadante G., Pannone E., Archidiacono N., Rocchi M.,
RA   Schlessinger D., D'Urso M.;
RT   "Differentially regulated and evolved genes in the fully sequenced
RT   Xq/Yq pseudoautosomal region.";
RL   Hum. Mol. Genet. 9:395-401(2000).
```

```
RN   [5]
RP   SEQUENCE FROM N.A.
RA   Rieder M.J., Armel T.Z., Carrington D.P., Chung M.-W., Lee K.L.,
RA   Poel C.L., Toth E.J., Yi Q., Nickerson D.A.;
RT   "SeattleSNPs. NHLBI HL66682 program for genomic applications, UW-
RT   FHCRC, Seattle, WA (URL: http://pga.gs.washington.edu).";
RL   Submitted (DEC-2001) to the EMBL/GenBank/DDBJ databases.
CC   -!- FUNCTION: This is a receptor for interleukin-9.
CC   -!- SUBCELLULAR LOCATION: Type I membrane protein and secreted.
CC   -!- DOMAIN: The WSXWS motif appears to be necessary for proper protein
CC       folding and thereby efficient intracellular transport and cell-
CC       surface receptor binding.
CC   -!- DOMAIN: The box 1 motif is required for JAK interaction and/or
CC       activation.
CC   -!- SIMILARITY: Belongs to the type I cytokine family of receptors.
CC       Subfamily 4.
CC   -!- SIMILARITY: Contains 1 fibronectin type III domain.
```
```
DR   EMBL; M84747; AAA58679.1; -.
DR   EMBL; S71404; AAB30844.2; ALT_SEQ.
DR   EMBL; S71420; AAD14081.1; -.
DR   EMBL; L39064; AAC29513.1; -.
DR   EMBL; AJ271736; CAB96817.1; -.
DR   EMBL; AY071830; AAL55435.1; -.
DR   PIR; B45268; B45268.
DR   Genew; HGNC:6030; IL9R.
DR   MIM; 300007; -.
DR   GO; GO:0005615; C:extracellular space; TAS.
DR   GO; GO:0005887; C:integral to plasma membrane; TAS.
DR   GO; GO:0004919; F:interleukin-9 receptor activity; TAS.
DR   GO; GO:0008283; P:cell proliferation; TAS.
DR   GO; GO:0007165; P:signal transduction; TAS.
DR   InterPro; IPR002996; Cytkn_recept_B/G.
DR   InterPro; IPR008957; FN_III-like.
DR   InterPro; IPR003531; Hemtprecepts_F1.
DR   PROSITE; PS01355; HEMATOPO_REC_S_F1; 1.
KW   Glycoprotein; Polymorphism; Receptor; Signal; Transmembrane.
FT   SIGNAL        1     40
FT   CHAIN        41    522       Interleukin-9 receptor.
FT   DOMAIN       41    270       Extracellular (Potential).
FT   TRANSMEM    271    291       Potential.
FT   DOMAIN      292    522       Cytoplasmic (Potential).
FT   DOMAIN      150    244       Fibronectin type-III.
FT   DOMAIN      429    439       Poly-Ser.
FT   DOMAIN      440    443       Poly-Asn.
FT   SITE        245    249       WSXWS motif.
FT   SITE        301    309       Box 1 motif.
FT   CARBOHYD    117    117       N-linked (GlcNAc...) (Potential).
FT   CARBOHYD    156    156       N-linked (GlcNAc...) (Potential).
FT   VARIANT     239    239       E -> Q (in dbSNP:6522).
FT                               /FTId=VAR_014804.
FT   CONFLICT    331    331       G -> R (in Ref. 1 and 2).
FT   CONFLICT    439    439       Missing (in Ref. 3 and 4).
SQ   SEQUENCE   522 AA;  57233 MW;  BBB73D6E2FAE37CB CRC64;

Query Match             9.2%;  Score 251;  DB 1;  Length 522;
Best Local Similarity  25.9%;  Pred No. 3.5e-09;
Matches  112;  Conservative  48;  Mismatches 164;  Indels 108;  Gaps 22;

Qy    45  RGPEE--LLCFTERLEDLVCFWEEAASAGVGPGNYSFSYQLEDEPWKLCRLHQAP-----  97
          ||  |     ||:     :|||  |    .: :: |  || :| :::|    ||||||:
Db    47  QGPRSRTFTCLTNNILRIDCHW-SAPELGQG----------SSPWLLFTSNQAPGGTHK  94
Qy    98  -TARGAVRFWCWCSLPTADTSSFVPLE-LRVT---AASG--------APRYHRVIHINE  141
          ||:  |  :    :      ||  ::    ||  :|   |         ||  || |
```

```
Db    95 CILRGSE---CTVVLPPEAVLVPSDNFTITFHHCMSGREQVSLVDPEYLPRRH------- 144
         : :||  | :||       ||
Qy   142 VVLLDAPVGLVARLADESGHVVLRW-LPPPETPMTSHIRYEVDVSAGNGA-GSVQRVEIL 199
          |||| |||  :: ||| |||| :: || ||| ||
Db   145 -VKLDPPSDLQSNIS--SGHCILTWSISPALEPMTTLLSYELAFKKQEEAWEQAQHRDHI 201
Qy   200 EGRTECVLSNLRGRTRYTFAVRAR--------MAEPSFGGFWSAWSEPVSLLTPSDLD 249
          ||    |    | :   |        ||   ||| :||||||||||||| ||
Db   202 VGVTWLLEAFELDPGFIHEARLRVQMATLEDDVVEERYTGQWSEWSQPVCFQAPQRQG 261
Qy   250 PLIILTL----SLILVVILVLLT----VLALLSHRRALKQKIWPGIPSPESEFEGLFTTH 300
          ||        :| || : :|        |:||  |
Db   262 PLIPEWGWPGNTLVAVSIFLLLTGPTYLLFKLSPR--VKRIFYQNVPSPAMFFQPLYSVH 319
Qy   301 KGNFQLWL-YQNDGCLWWSPCT--------PFTEDPPASLEVLSERCWGTMQAVEPGTDD 351
          :|||  :  |  | ||  :          |:|:||| |   || ::
Db   320 NGNFQTWMGAHGAVGAGVLLSQDCAGTPQGALEPCVQEATALLTCGPARPWKSVALEE---EQ 376
Qy   352 EGPLLEPVGSEHAQD------------TYLVLDKWLLPRNPPSEDLPGPGGSVDIVAM 397
          ||  |           :::||
Db   377 EGPGTRLPGNLSSEDVLPAGCTEWRVQTLAYLPQEDWA----PTSLTRPAPPDS------ 426
Qy   398 DEGSEASSCCSSA 409
          :|| || |
Db   427 -EGSRSSSSSSS 437

RESULT 12
Q6IYE8
ID   Q6IYE8      PRELIMINARY;    PRT;   630 AA.
AC   Q6IYE8;
DT   05-JUL-2004 (TrEMBLrel. 27, Created)
DT   05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT   05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE   Thrombopoietin receptor.
GN   Name=mpl;
OS   Gallus gallus (Chicken).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC   Gallus.
OX   NCBI_TaxID=9031;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Bartunek P., Karafiat V., Bartunkova J., Dvorakova M., Kralova J.,
RA   Pajer P., Zenke M., Dvorak M.;
RL   Submitted (APR-2004) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AY613435; AAT45555.1; -.
DR   GO; GO:0004872; F:receptor activity; IEA.
DR   InterPro; IPR003961; FN_III.
DR   InterPro; IPR008957; FN_III-like.
DR   Pfam; PF00041; fn3; 2.
DR   SMART; SM00060; FN3; 2.
DR   PROSITE; PS50853; FN3; 1.
KW   Receptor.
SQ   SEQUENCE 630 AA;  68856 MW;  9B8FCC6CE61E0EDE CRC64;

Query Match    8.5%;        Score 231.5;  DB 2;  Length 630;
Best Local Similarity   26.6%;  Pred. No. 1e-07;
Matches  71;  Conservative   32;  Mismatches  107;  Indels   57;  Gaps   9;

Qy   24 WAP----------PPNLPDPKFESKAALLAARGPEELLCFTERLEDLVCFWEEAAASAG 71
         | |          |||: || |||:|| | ::  |||:|||| |||||:|||||: :|
Db   13 WLPALLTAVLLSRHSPATDTEPVTSQDAAIL-AEVPEDILCFSRSFEDLTCFWDEEEASA 71
Qy   72 VGPGNYSFSYQLEDEPWKLCRLHQAPTARGAVRFWCSLPTADTSSFVPLELRV-TAASGA 130
         :    :  | :| |    ||| :  : |    :  |  |
Db   72 I----HRFYYW--DAP-TTCAVSTQRSGVSGMRHICVFPSQDVRLFTPLHVLVLDTTTNR 124
Qy  131 PRYHVIHINEVVLLDAPVGLVARLADESGHVLRWLPP---------------PETPMT 175
          |      ||:||||||||||||||||||| |||:||               |||||
Db  125 TRHRRELSVDAVGLIAPPVNISARWAGAAGQLRVSWQPPLADYPNFFLYEVCVHPMTPTV 184
Qy  176 --------------SHIRYEVDVSAGNGAGSVQRVEILEGRTECVLSNLRGRTRYTF 218
```

```
Db  185 TPRGTAPGDLPARPTGRAHPPTARGAAASQGTGQ---GLVQADTWVLRELQPGVRYHI 240
         :         | | | ||            |     ||:||||||| ||||| | |
Qy  219 AVRARMAEPSFGGFWSAWSEPVSLLTP 245
          :|:  || ||| : ||: || ||
Db  241 QVRSKPDGTSMDGVWGPWSEVVVAETP 267

RESULT 13
Q925F5
ID   Q925F5      PRELIMINARY;    PRT;   566 AA.
AC   Q925F5;
DT   01-DEC-2001 (TrEMBLrel. 19, Created)
DT   01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT   01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE   C-mpl-II.
GN   Name=mpl;
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=21992449; PubMed=11997107; DOI=10.1016/S0167-4781(01)00357-8;
RA   Sabath D.F., Lofton-Day C., Lin N., Lok S., Kenneth Kaushansky.;
RA   Broudy V.C.;
RT   "Identification and characterization of an isoform of murine Mpl.";
RL   Biochim. Biophys. Acta 1574:383-386(2002).
DR   EMBL; AF360122; AAK52492.1; -.
DR   HSSP; P19235; 1ERN.
DR   MGD; MGI:97076; Mpl.
DR   GO; GO:0005615; C:extracellular space; TAS.
DR   GO; GO:0016021; C:integral to membrane; TAS.
DR   Pfam; PF00041; fn3; 1.
DR   SMART; SM00060; FN3; 1.
DR   PROSITE; PS50853; FN3; 1.
DR   PROSITE; PS01352; HEMATOPO_REC_L_F1; UNKNOWN_1.
DR   PROSITE; PS00041; HTH_ARAC_FAMILY_1; UNKNOWN_1.
SQ   SEQUENCE 566 AA;  63339 MW;  A19D3DD5F7BDBA6B CRC64;

Query Match    8.3%;   Score 224;  DB 2;  Length 566;
Best Local Similarity   21.8%;  Pred. No. 3e-07;
Matches  128;  Conservative   42;  Mismatches  167;  Indels  250;  Gaps   23;

Qy   16 CLLLAGAAWAPPPNLPD-PKFESKAALLAARGPEELLCFTERLEDLVCFWEEAAASAGVGP 74
         |||||     || |      |:: |||||| |||| || || |||| |||:||||    ||
Db    5 CLLLA------LPNQAQVTSQDVFLLALGTEPLNCFSQTFEDLTCFWDEEEA--PS 53
Qy   75 GNYSFSYQLEDEPWKLCRLHQAPTARGAVRFWCSLPTAD-TSSFVPLELRVTAAS-GAPR 132
         | | | |   |:   |   :   : |    | |:  |:|  : |   |   :| |  || |
Db   54 GTYQLLYAYRGEKPRACPLYSQSVPTFGTRYVCQFPAQDEVRLFFPLHLWYKNVSLNQTL 113
Qy  133 YHRVIHI---------------NEVVLLDAP-------------- 148
          :|| :|                |:|:| ||
Db  114 IQRVLFVDSVETCCPTLWMPNPVPVLDQPPCVHPTASQPHGPVRTSPAGEAPFLTVKGGS 173
Qy  149 ---VGLVA-------RLADESGHVVLR--WLP-------------- 168
            |||:        || :: | | |:   ||:
Db  174 CLVSGLQAGKSYWLQLRSQPDGVSLRGSWGPWSFPVTVDLPGDAVTIGLQCFTLDLKMVT 233
Qy  169 --------------PPETPM---------TSH 177
                        ||||:            |:|
Db  234 CQWQQQDRTSSQGFFRHSRTRCCPTDRDPTWEKCEEEPRPGSQPALVSRCHFKSRNDSV 293
Qy  178 IRYEVDVSAGNGA---------GSVQRVEILEGRTE-------- 204
         | : | :: :|           ||::||| |
Db  294 IHILVEVTTAQGAVHSYLGSPFWIHQAVLLPTPSLHWREVSSGRLELEWQHQSSWAAQET 353
Qy  205 C-------------------VLS---------------NLRGRTRYTFAVRARMAEPSFGGFWSAWSEPV 240
         |                   ||:               ||:| ||||| | |||||| |||||||| ||
Db  354 CYQLRYTGEGREDWKVLEPSLGARGGTLELRPRARYSLQLRARLNGPTYQGPWSAWSPPA 413
Qy  241 SLLTPSDLDPLILTLSLILVVILVLLTVLALL----SHRRALKQKIWPGIPSPESEFEG 295
```

```
Db   414  RVSTGSETAWITILVTALLVLSLSSALLGLLLLKWQFPAHYRRLRHALWPSLPD------  466
              :  | |:   :||:   |  | |     || |    | :|::||| | :  ||  :|
Qy   296  LFTTHKGNFQLWLYQNDGCLWWSPCTPFTED-----PPASLEVLSERCWGT---MQAVEP  347
              |:    |:|||          |:||| | |       :|:||| | | |
Db   467  ---LHR---VLGQYLRDTAA-LSPSKATVTDSCEEVEPSLLEILPKSSESTPLPLCPSQP  519

Qy   348  GTDDEG---------PLLEPVGS---EHAQDTYLVLDKWLLP                    377
             |                  |:|        :||  |: |||
Db   520  QMDYRGLQPCLRTMPLSVCPPMAETGSCCTTHIANHSYLPLSYWQQP              566
```

RESULT 14
PRLR_PIG
ID   PRLR_PIG   STANDARD;   PRT;   625 AA.
AC   Q6JTA8; O46679; Q6JTA9; Q6JTB0;
DT   25-OCT-2004 (Rel. 45, Created)
DT   25-OCT-2004 (Rel. 45, Last sequence update)
DT   25-OCT-2004 (Rel. 45, Last annotation update)
DE   Prolactin receptor precursor (PRL-R).
GN   Name=PRLR;
OS   Sus scrofa (Pig).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
OX   NCBI_TaxID=9823;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Landrace, and Yucatan;
RA   Trott J.F., Hovey R.C.;
RT   "Cloning and characterization of multiple forms of the pig (Sus
RT   scrofa) prolactin receptor.";
RL   Submitted (MAY-2003) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   SEQUENCE OF 531-625 FROM N.A.
RX   MEDLINE=97468290; PubMed=9321486;
RA   Vincent A.L., Wang L., Tuggle C.K., Robic A., Rothschild M.F.;
RT   "Prolactin receptor maps to pig chromosome 16.";
RL   Mamm. Genome 8:793-794(1997).
CC   -!- FUNCTION: This is a receptor for the anterior pituitary hormone
CC   prolactin (By similarity).
CC   -!- SUBCELLULAR LOCATION: Type I membrane protein (By similarity).
CC   -!- DOMAIN: The WSXWS motif appears to be necessary for proper protein
CC   folding and thereby efficient intracellular transport and cell-
CC   surface receptor binding (By similarity).
CC   -!- DOMAIN: The box 1 motif is required for JAK interaction and/or
CC   activation.
CC   -!- SIMILARITY: Belongs to the type I cytokine family of receptors.
CC   Subfamily 1.
CC   -!- SIMILARITY: Contains 2 fibronectin type III domains.
CC   ---
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---
DR   EMBL; AY308824; AAQ76841.1; -.
DR   EMBL; AY308825; AAQ76842.1; -.
DR   EMBL; AY308826; AAQ76843.1; -.
DR   EMBL; U96310; AAC02897.1; -.
DR   InterPro; IPR002996; Cytkn_recept_B/G.
DR   InterPro; IPR003961; FN_III.
DR   InterPro; IPR008957; FN_III-like.
DR   InterPro; IPR003528; HemreceptL_F1.
DR   Pfam; PF00041; fn3; 2.
DR   SMART; SM00060; FN3; 2.
DR   PROSITE; PS50853; FN3; 2.
DR   PROSITE; PS01352; HEMATOPO_REC_L_F1; 1.
KW   Glycoprotein; Receptor; Repeat; Signal; Transmembrane.
FT   SIGNAL   1   24   Potential.
FT   CHAIN   25   625   Prolactin receptor.

FT   DOMAIN   25   237   Extracellular (Potential).
FT   TRANSMEM   238   259   Potential.
FT   DOMAIN   259   625   Cytoplasmic (Potential).
FT   DOMAIN   27   123   Fibronectin type-III 1.
FT   DOMAIN   127   227   Fibronectin type-III 2.
FT   DOMAIN   111   114   Poly-Ser.
FT   SITE   215   219   WSXWS motif.
FT   SITE   267   275   Box 1 motif.
FT   DISULFID   36   46   By similarity.
FT   DISULFID   75   86   By similarity.
FT   CARBOHYD   59   59   N-linked (GlcNAc...) (Potential).
FT   CARBOHYD   104   104   N-linked (GlcNAc...) (Potential).
FT   CARBOHYD   132   132   N-linked (GlcNAc...) (Potential).
FT   CONFLICT   372   372   P -> R (in Ref. 1; AAQ76842).
FT   CONFLICT   406   406   L -> R (in Ref. 1; AAQ76842).
FT   CONFLICT   461   461   A -> G (in Ref. 1; AAQ76842).
FT   CONFLICT   480   480   K -> R (in Ref. 1; AAQ76842).
FT   CONFLICT   510   510   M -> L (in Ref. 1; AAQ76842).
FT   CONFLICT   534   534   G -> S (in Ref. 1; AAQ76842).
FT   CONFLICT   597   597   S -> G (in Ref. 1; AAQ76843).
SQ   SEQUENCE   625 AA; 69572 MW; 751B2B593ABFC51B CRC64;

Query Match   8.1%; Score 220; DB 1; Length 625;
Best Local Similarity   23.8%;
Matches 113; Conservative   65; Mismatches 179; Indels 117; Gaps 25;
```
Qy    27  PPNLPDPKFESKAALLAARGPEELLCFTERLEDLVCFWEEAASAGVGPGNYSFSYQLEDE   86
             |  |:                 | :                    :  |:  | :|::
Db    27  PPGKPE------------IFKCRSPEK--------ETFTCWWKPGADGGL-PTNYLTYHKEGE  69

Qy    87  PWKLCRLHQAPTARGAVRFWCSLPTADTSSFVPLELRVTA------ASGAPRYHRVIHIN  140
              |  :      :        | :             |: | :       | ||:  |:
Db    70  TF----THECPDYKTGGPNSCYFNKGHTSIWTIYIITVNATNQMGSSSSDPRY---VDVT  122

Qy   141  EVVLLDAPVGLVARL---ADESGHVVLRWLPPPETPMTS---HIRYEVDVSAGNGAGSVQ  194
               |: |||  | ||      |  :  | |||||   ||     | |:|||:|| :
Db   123  YIVEPDPPVNLTLELKKPEDQKPYLWIKWLPPTLVDVRSGWLTLQYEIRLQPEKTA--EW  180

Qy   195  RVEILEGRTECVLSNLRGRTRYTFAVRARMAEPSFGGFWSAWSEPVSLLTPSDLDPLLT   254
            :  | | | | |   | | : :  |   |   | |||| |||||:::  |:  || ||
Db   181  ETHFAGGQTQFKILSLYPGGQKYLVQVRCK---PDH-GFWSEWSPESSIQIFNDFSMKDTT  236

Qy   255  LSLIL-----VVILVLLTVLALLSHRRALKQKIWPGIPSPESEFEGLFTTH---KGNFQL  306
             :: |      | |  |:::|:|  |:     :     |:    | |:|     |    |
Db   237  MWIFVAILSAVVCLIMIWAVALKGY--SMVACILPPVGP--KIKG-FDTHLLEKGK---  288

Qy   307  WLYQNDGCLWWSPCTPFTEDPPASLEVLSE-RCWGTMQAVEPGTDDEGPLLEPVGSEHAQ  365
                          :  | |: |        :   |
Db   289  -----------SEELLSALGC----QDFPPTSDCEDLLVEFLEVDDSE             321

Qy   366  DTYLVLDKWLLPRNPPSEDLPGPGGSVDIVAMDEGSEASSCSS---------ALASKP-  414
              : |:    | |      :     |      |: :           :   |
Db   322  ----DQQLMPAH--SKEHPSQGRKPTHLDPDSDSGRGSCCDPSLLSEKCDEPRANPPK   373

Qy   415  --SPEGASAASFEYTILDPSSQLLRPWTLCPELPPTPHLKYLYLVVSDSGIST        466
                |||:|                       : | | |   |||:|| | :|:|
Db   374  FHTPEGIEKPG------DPETNLPRP----QDPQSTSVESKLLYFHADGSKSST       417
```

RESULT 15
PRLR_CEREL
ID   PRLR_CEREL   STANDARD;   PRT;   581 AA.
AC   Q28235;
DT   01-NOV-1997 (Rel. 35, Created)
DT   01-NOV-1997 (Rel. 35, Last sequence update)
DT   05-JUL-2004 (Rel. 44, Last annotation update)
DE   Prolactin receptor precursor (PRL-R).
GN   Name=PRLR;
OS   Cervus elaphus (Red deer).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Cervidae;
OC   Cervinae; Cervus.
OX   NCBI_TaxID=9860;
```

```
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Liver;
RX   MEDLINE=96030711; PubMed=7561644;
RA   Clarke L.A., Edery M., Loudon A.S., Randall V.A., Postel-Vinay M.C.,
RA   Kelly P.A., Jabbour H.N.;
RT   "Expression of the prolactin receptor gene during the breeding and
RT   non-breeding seasons in red deer (Cervus elaphus): evidence for the
RT   expression of two forms in the testis.";
RL   J. Endocrinol. 146:313-321(1995).
CC   -!- FUNCTION: This is a receptor for the anterior pituitary hormone
CC       prolactin.
CC   -!- SUBCELLULAR LOCATION: Type I membrane protein.
CC   -!- DOMAIN: The WSXWS motif appears to be necessary for proper protein
CC       folding and thereby efficient intracellular transport and cell-
CC       surface receptor binding.
CC   -!- DOMAIN: The box 1 motif is required for JAK interaction and/or
CC       activation.
CC   -!- SIMILARITY: Belongs to the type I cytokine family of receptors.
CC       Subfamily 1.
CC   -!- SIMILARITY: Contains 2 fibronectin type III domains.
CC   ----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ----------------------------------------------------------------------
DR   EMBL; X94953; CAA64419.1; -.
DR   HSSP; P05710; 1P6F.
DR   GO; GO:0009986; C:cell surface; ISS.
DR   GO; GO:0042978; F:ornithine decarboxylase activator activity; ISS.
DR   GO; GO:0004925; F:prolactin receptor activity; ISS.
DR   GO; GO:0042803; F:protein homodimerization activity; ISS.
DR   GO; GO:0006916; P:anti-apoptosis; ISS.
DR   GO; GO:0007166; P:cell surface receptor linked signal transdu. . .; ISS.
DR   GO; GO:0007595; P:lactation; ISS.
DR   GO; GO:0006694; P:steroid biosynthesis; ISS.
DR   GO; GO:0042110; P:T-cell activation; ISS.
DR   GO; GO:0007171; P:transmembrane receptor protein tyrosine kin. . .; ISS.
DR   GO; GO:0042977; P:tyrosine phosphorylation of JAK2 protein; ISS.
DR   InterPro; IPR002996; Cytkn_recept_B/G.
DR   InterPro; IPR003961; FN_III.
DR   InterPro; IPR008957; FN_III-like.
DR   InterPro; IPR003528; Hemptrecept L_F1.
DR   Pfam; PF00041; fn3; 2.
DR   PROSITE; PS00853; FN3; 2.
DR   PROSITE; PS01352; HEMATOPO_REC_L_F1; 1.
KW   Glycoprotein; Receptor; Repeat; Signal; Transmembrane.
FT   SIGNAL        1     24       By similarity.
FT   CHAIN        25    581       Prolactin receptor.
FT   DOMAIN       25    234       Extracellular (Potential).
FT   TRANSMEM    235    258       Potential.
FT   DOMAIN      259    581       Cytoplasmic (Potential).
FT   DOMAIN       27    123       Fibronectin type-III 1.
FT   DOMAIN      127    227       Fibronectin type-III 2.
FT   SITE        215    219       WSXWS motif.
FT   SITE        267    275       Box 1 motif.
FT   DISULFID     36     46       By similarity.
FT   DISULFID     75     86       By similarity.
FT   CARBOHYD     59     59       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD    132    132       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD    233    233       N-linked (GlcNAc. . .) (Potential).
SQ   SEQUENCE   581 AA;  65159 MW;  975E47CB63CF28EC CRC64;
```

Query Match 8.0%; Score 216.5; DB 1; Length 581;
Best Local Similarity 21.6%; Pred. No. 1e-06;
Matches 113; Conservative 80; Mismatches 206; Indels 125; Gaps 22;

```
Qy   47 PEELLCFTERLEDLVCFWEEAASAGVGPGNYSFSYQLEDEPWKLCRLHQAPTARGAVRFW 106

Db   31 PKIIKCRSPGKETTCWWEPGSDGGL-PTNYTLTYHKEGE----TLIHECPDYKTGGPNT 85
Qy  107 CSLPTADTSSFVPLELRVTA------ASGAPRYHRVIHINEVVLLDAPVGLVARL---AD 157
Db   86 CYFSKKHTSIWKIYVITVNAINQMGVSSSDPLY---VDVTYIVEPEPPANLTELKHPED 142
Qy  158 ESGHVLRWLPPPETPMTS--HIRYEVDVSAGNGAGSVQRVEILEGRTECVLSNLRGRT 214
Db  143 RKPYLWIKWFPPTLTDVKSGWFMIQYEIRLKPETAAD--WEIHFAAKQTQLKIFSLYPGQ 200
Qy  215 RYTFAVRARMAEPSFGGFWSAWSEPVSLLTPSDL---DPLILTLSLILVVILVLLTVLAL 271
Db  201 KYLVQVRCK---PDH-GYWSEWSPESSIQIPNDFPVNDTTVWIFVAVLSAVICLIMVWAV 256
Qy  272 LSHRRALKQKIWPGIPSPESEFEGLFTTHKGNFQLWLYQNDGCLWWSPCTPFTEDPPASL 331
Db  257 ALKGYSMMTCILPPVPGPKIKGFDIHLLEKGK-------------SEELLRAL 296
Qy  332 EVLSERCWGTMQAVEPGTDDEGPLLEPVGSEHAQDTYLVLDKWLLPRNPPSEDLPGPGGS 391
Db  297 E------SQDFPPTSDCEDLLMEFIEVDDSE------DQQLMPR--PSKEHMEQGVK 339
Qy  392 VDIVAMDEGSEASSCSS-ALASKPSPEGASAASFEYTILDPSSQLLRPWT--LCPELP-- 446
Db  340 PMHMDPDSDSGRGSCDSPSLFSEKCDEPQAHPFKFYTPEDP-BKLENPETNLTCLQAPQS 398
Qy  447 ----------------PTPPHL---KYLYLVSD-----------SGIS--------- 465
Db  399 TSREDKIPYFHANGPKSSTWPPFQPPSLHNPRYSYHNIADVCELALGMAGTTATLLDQTD 458
Qy  466 --------TDYSSGDSQGGAQGGLSDGPYSNPYENSLIPAAEPLP 501
Db  459 QHALKPSKTIETGGEGKAAKQRESEGCSSKPDQDT----GWPLP 498
```

Search completed: April 21, 2005, 08:59:16
Job time : 221.5 secs

```
                    GenCore version 5.1.6
         Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on:      April 21, 2005, 04:08:54 ; Search time 189 Seconds
                                          (without alignments)
                                    1039.546 Million cell updates/sec

Title:       US-09-016-159D-5
Perfect score: 2715
Sequence:    1 MDHLGASLWPQVGSLCLLLA.........YENSLIPAAEPLPPSYVACS 508
Scoring table: BLOSUM62
             Gapop 10.0 , Gapext 0.5
Searched:    2105692 seqs, 386760381 residues

Total number of hits satisfying chosen parameters:   2105692

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :   A_Geneseq_16Dec04:*
             1: geneseqp1980s:*
             2: geneseqp1990s:*
             3: geneseqp2000s:*
             4: geneseqp2001s:*
             5: geneseqp2002s:*
             6: geneseqp2003as:*
             7: geneseqp2003bs:*
             8: geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

                        SUMMARIES
```

Result No.	Score	% Query Match	Score Length	DB	ID	Description
1	2715	100.0	508	2	AAR06512	Aar06512 EPO recep
2	2715	100.0	508	2	AAR70032	Aar70032 Human ery
3	2715	100.0	508	2	AAR69503	Aar69503 Human ery
4	2715	100.0	508	8	ADO05724	Ado05724 Human ery
5	2715	100.0	508	8	ADP12417	Adp12417 Protein e
6	2715	100.0	508	8	ABM82037	Abm82037 Tumour-as
7	2708	99.7	508	5	ABB09173	Abb09173 Human ery
8	2693	99.2	508	8	ADM93422	Adm93422 Human NOV
9	2691	99.1	508	8	ADE28677	Ade28677 Human NOV
10	2690	99.1	508	2	AAR47518	Aar47518 Human EPO
11	2407	88.7	458	7	ADE28675	Ade28675 Human NOV
12	2407	88.7	458	8	ADM93420	Adm93420 Human NOV
13	2404	88.5	458	7	ADE28671	Ade28671 Human NOV
14	2404	88.5	458	7	ADE28673	Ade28673 Human NOV
15	2404	88.5	458	8	ADM93416	Adm93416 Human NOV
16	2404	88.5	458	8	ADM93418	Adm93418 Human NOV
17	2330	85.8	438	3	AAY44622	Aay44622 Truncated
18	2322	85.5	438	3	AAY44623	Aay44623 R154C tru
19	2171.5	80.0	507	2	AAR06511	Aar06511 EPO recep
20	2169.5	79.9	507	2	AAR47517	Aar47517 MEL EPO r
21	2169.5	79.9	507	2	AAR69502	Aar69502 Mouse ery
22	2167.5	79.8	507	2	AAR50327	Aar50327 Mouse sol
23	1352.5	49.8	503	3	AAB13012	Aab13012 Q-tagged
24	1335	49.2	488	2	AAW08349	Aaw08349 EporRFc fu
25	1194	44.0	225	3	AAB21685	Aab21685 Human mat

26	1112	41.0		3	AAB21686	Aab21686 Human mat
27	1063.5	39.2		2	AAR50326	Aar50326 Mouse sol
28	371.5	13.7		8	ADL92633	Adl92633 IMABIS019
29	370.5	13.6		8	ADL92637	Adl92637 IMABIS028
30	370.5	13.6		8	ADR41983	Adr41983 VAP relat
31	370.5	13.6		8	ADR41979	Adr41979 VAP relat
32	366	13.5		8	ADL92635	Adl92635 IMABIS025
33	366	13.5		8	ADR41981	Adr41981 VAP relat
34	354.5	13.1		3	AAY94338	Aay94338 Human cel
35	332.5	12.2		8	ADL92634	Adl92634 IMABIS020
36	332.5	12.2		8	ADR41980	Adr41980 VAP relat
37	325.5	12.0		8	ADL92636	Adl92636 IMABIS027
38	325.5	12.0		8	ADR41982	Adr41982 VAP relat
39	286.5	10.6		8	ADL92535	Adl92535 iMabl100
40	286.5	10.6		8	ADR41849	Adr41849 Versatile
41	275.5	10.1		8	ADL92527	Adl92527 iMab600 V
42	275.5	10.1		8	ADR41842	Adr41842 Versatile
43	261.5	9.6		2	AAW64057	Aaw64057 Human IL-
44	251	9.2		2	AAW64055	Aaw64055 Human IL-
45	251	9.2		6	AAO19700	Aao19700 IL-13 mod

```
                            ALIGNMENTS

RESULT 1
AAR06512
ID   AAR06512 standard; protein; 508 AA.
XX
AC   AAR06512;
XX
DT   25-MAR-2003  (revised)
DT   04-JAN-1991  (first entry)
XX
DE   EPO receptor.
XX
KW   Erythropoietin; Diamond Blackfan anaemia; polycythemia vera.
XX
OS   Homo sapiens.
XX
PN   WO9008822-A.
XX
PD   09-AUG-1990.
XX
PF   03-FEB-1989;   89US-00306503.
XX
PR   03-FEB-1989;   89US-00306503.
XX
PA   (GEMY ) GENETICS INST INC.
PA   (WHED ) WHITEHEAD INST BIOMEDICAL RES.
XX
PI   Dandrea A, Wong G;
XX
DR   WPI; 1990-260931/34.
DR   N-PSDB; AAQ05748.
XX
PT   Erythropoietin receptor and gene - used for developing reagents and
PT   systems to control and study erythropoiesis.
XX
PS   Disclosure; Fig 2; 53pp; English.
XX
CC   The sequence was deduced from DNA obtd. from a clone isolated from a
CC   commercially available human genomic cDNA library in phage Lambda Fix
CC   (Stratagene). The sequence encodes a type I trans- membrane protein with
CC   binding affinity for EPO. The gene and recombinant EPO recepor produced
CC   on expression of the DNA are used to develop reagents and systems to
CC   control and study erythropoiesis. It is believed that the EPO receptor is
CC   dys- functional in individuals with Diamond Blackfan anaemia, and may be
CC   hyperactive in polycythemia vera. See also AAR06511 (murine EPO
CC   receptor). (Updated on 25-MAR-2003 to correct PA field.) (Updated on 25-
CC   MAR-2003 to correct PI field.)
XX
SQ   Sequence 508 AA;
```

RESULT 2
AAR70032

Query Match   100.0%; Score 2715; DB 2; Length 508;
Best Local Similarity 100.0%; Pred. No. 1.9e-216;
Matches 508; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy    1 MDHLGASLWPQVGSLCLLLAGAAWAPPNLPDPKFESKAALLAARGPEELLCFTERLEDL  60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1 MDHLGASLWPQVGSLCLLLAGAAWAPPNLPDPKFESKAALLAARGPEELLCFTERLEDL  60

Qy   61 VCFWEEAASAGVGPGNYSFSYQLEDEPWKLCRLHQAPTARGAVRFWCSLPTADTSSFVPL 120
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61 VCFWEEAASAGVGPGNYSFSYQLEDEPWKLCRLHQAPTARGAVRFWCSLPTADTSSFVPL 120

Qy  121 ELRVTAASGAPRYHRVIHINEVVLLDAPVGLVARLADESGHVVLRWLPPPETPMTSHIRY 180
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  121 ELRVTAASGAPRYHRVIHINEVVLLDAPVGLVARLADESGHVVLRWLPPPETPMTSHIRY 180

Qy  181 EVDVSAGNGAGSVQRVEILEGRTECVLSNLRGRTRYTFAVRARMAEPSFGGFWSAWSEPV 240
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  181 EVDVSAGNGAGSVQRVEILEGRTECVLSNLRGRTRYTFAVRARMAEPSFGGFWSAWSEPV 240

Qy  241 SLLTPSDLDPLILTLSLILVVILVILVLLTVLALLSHRRALKQKIWPGIPSPESEFEGLFTTH 300
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  241 SLLTPSDLDPLILTLSLILVVILVILVLLTVLALLSHRRALKQKIWPGIPSPESEFEGLFTTH 300

Qy  301 KGNFQLWLYQNDGCLWWSPCTPFTEDDPPASLEVLSERCWGTMQAVEPGTDDEGPLLEPVG 360
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  301 KGNFQLWLYQNDGCLWWSPCTPFTEDDPPASLEVLSERCWGTMQAVEPGTDDEGPLLEPVG 360

Qy  361 SEHAQDTYLVLDKWLLPRNPPSEDLPGPGGSVDIVAMDEGSEASSCSSALASKPSPEGAS 420
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  361 SEHAQDTYLVLDKWLLPRNPPSEDLPGPGGSVDIVAMDEGSEASSCSSALASKPSPEGAS 420

Qy  421 AASFEYTILDPSSQLLRPWTLCPELPPTPPHLKYLYLVVSDSGISTDYSSGDSQGAQGGL 480
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  421 AASFEYTILDPSSQLLRPWTLCPELPPTPPHLKYLYLVVSDSGISTDYSSGDSQGAQGGL 480

Qy  481 SDGPYSNPYENSLIPAAEPLPPSYVACS 508
          ||||||||||||||||||||||||||||
Db  481 SDGPYSNPYENSLIPAAEPLPPSYVACS 508
```

RESULT 2
AAR70032
ID   AAR70032 standard; protein; 508 AA.
XX
AC   AAR70032;
XX
DT   25-MAR-2003 (revised)
DT   07-OCT-1995 (first entry)
XX
DE   Human erythropoietin receptor.
XX
KW   Erythropoietin receptor; extracellular domain.
XX
OS   Homo sapiens.
XX
FH   Key   Location/Qualifiers
FT   Domain   9..83
FT   /note= "extracellular domain"
FT   Domain   25..250
FT   /note= "extracellular domain"
FT   Misc-difference   25..29
FT   /note= "forward primer AAQ82991 specific site"
FT   Misc-difference   222..226
FT   /note= "reverse primer AAQ82992 specific site"
XX
PN   WO9505469-A1.
XX
PD   23-FEB-1995.
XX
PF   15-AUG-1994; 94WO-US009298.
XX

PR   16-AUG-1993; 93US-00106815.
XX
PA   (LEEJ/) LEE J Y.
XX
PI   Lee JY;
XX
DR   WPI; 1995-098767/13.
DR   N-PSDB; AAQ82990.
XX
PT   New pure human erythropoietin receptor fragment - obtd. by expression as
PT   a fusion protein having a thrombin proteolytic cleavage site.
XX
PS   Disclosure; Page 27-29; 42pp; English.
XX
CC   The full-length erythropoietin receptor (EPO-R) is given. Extracellular
CC   domains are expressed from vector plasmid pGEX-2T as fusion proteins with
CC   glutathione-S-transferase. The domains are used for invesigating the
CC   structure of the EPO-R and for identifying factors involved in regulating
CC   differentiation and proliferation mechanisms in erythroid progenitor
CC   cells. They can also be used for identifying and quantitating EPO and EPO
CC   -R as well as in understanding haematopoietic malignancy and some
CC   cardiovascular system disorders. (Updated on 25-MAR-2003 to correct PN
CC   field.)
XX
SQ   Sequence 508 AA;

Query Match   100.0%; Score 2715; DB 2; Length 508;
Best Local Similarity 100.0%; Pred. No. 1.9e-216;
Matches 508; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy    1 MDHLGASLWPQVGSLCLLLAGAAWAPPNLPDPKFESKAALLAARGPEELLCFTERLEDL  60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1 MDHLGASLWPQVGSLCLLLAGAAWAPPNLPDPKFESKAALLAARGPEELLCFTERLEDL  60

Qy   61 VCFWEEAASAGVGPGNYSFSYQLEDEPWKLCRLHQAPTARGAVRFWCSLPTADTSSFVPL 120
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61 VCFWEEAASAGVGPGNYSFSYQLEDEPWKLCRLHQAPTARGAVRFWCSLPTADTSSFVPL 120

Qy  121 ELRVTAASGAPRYHRVIHINEVVLLDAPVGLVARLADESGHVVLRWLPPPETPMTSHIRY 180
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  121 ELRVTAASGAPRYHRVIHINEVVLLDAPVGLVARLADESGHVVLRWLPPPETPMTSHIRY 180

Qy  181 EVDVSAGNGAGSVQRVEILEGRTECVLSNLRGRTRYTFAVRARMAEPSFGGFWSAWSEPV 240
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  181 EVDVSAGNGAGSVQRVEILEGRTECVLSNLRGRTRYTFAVRARMAEPSFGGFWSAWSEPV 240

Qy  241 SLLTPSDLDPLILTLSLILVVILVILVLLTVLALLSHRRALKQKIWPGIPSPESEFEGLFTTH 300
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  241 SLLTPSDLDPLILTLSLILVVILVILVLLTVLALLSHRRALKQKIWPGIPSPESEFEGLFTTH 300

Qy  301 KGNFQLWLYQNDGCLWWSPCTPFTEDDPPASLEVLSERCWGTMQAVEPGTDDEGPLLEPVG 360
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  301 KGNFQLWLYQNDGCLWWSPCTPFTEDDPPASLEVLSERCWGTMQAVEPGTDDEGPLLEPVG 360

Qy  361 SEHAQDTYLVLDKWLLPRNPPSEDLPGPGGSVDIVAMDEGSEASSCSSALASKPSPEGAS 420
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  361 SEHAQDTYLVLDKWLLPRNPPSEDLPGPGGSVDIVAMDEGSEASSCSSALASKPSPEGAS 420

Qy  421 AASFEYTILDPSSQLLRPWTLCPELPPTPPHLKYLYLVVSDSGISTDYSSGDSQGAQGGL 480
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  421 AASFEYTILDPSSQLLRPWTLCPELPPTPPHLKYLYLVVSDSGISTDYSSGDSQGAQGGL 480

Qy  481 SDGPYSNPYENSLIPAAEPLPPSYVACS 508
          ||||||||||||||||||||||||||||
Db  481 SDGPYSNPYENSLIPAAEPLPPSYVACS 508
```

RESULT 3
AAR69503
ID   AAR69503 standard; protein; 508 AA.
XX
AC   AAR69503;
XX

```
DT   25-MAR-2003 (revised)
DT   11-AUG-1995 (first entry)
XX
DE   Human erythropoietin receptor.
XX
KW   Erythropoietin receptor; anemia therapy; diagnostic.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   Peptide         1..24
FT                   /note= "signal peptide"
FT   Protein         25..508
FT                   /note= "mature protein"
FT   Modified-site   76..79
FT                   /note= "N-glycosylation site"
FT   Domain          251..272
FT                   /note= "transmembrane region"
XX
PN   US5378808-A.
XX
PD   03-JAN-1995.
XX
PF   10-JUN-1993; 93US-00075069.
XX
PR   03-FEB-1989; 89US-00306503.
PR   25-MAR-1991; 91US-00678877.
XX
PA   (GEMY ) GENETICS INST INC.
XX
PI   Jones SS, D'andrea A, Wong GG;
XX
DR   WPI; 1995-051310/07.
DR   N-PSDB; AAQ81892.
XX
PT   recombinant erythropoietin receptor polypeptide(s) - used for detection,
PT   purificn. and therapy and for prodn. of antibodies for detection and
PT   therapy.
XX
PS   Claim 2; Fig 9; 24pp; English.
XX
CC   The sequence is that of a 55-kDa human erythropoietin receptor. The
CC   receptor polypeptide may be used in purification and detection of
CC   erythropoietin, and in production of antibodies for anemia therapy. The
CC   polypeptide may also be used for treating individuals who are
CC   hypersensitive to erythropoietin or who have elevated erythropoietin
CC   levels. They may be used in therapy of e.g. primary or secondary
CC   proliferative polycythemia. (Updated on 25-MAR-2003 to correct PF field.)
XX
SQ   Sequence 508 AA;

  Query Match             100.0%;  Score 2715;  DB 2;  Length 508;
  Best Local Similarity   100.0%;  Pred. No. 1.9e-216;
  Matches  508;  Conservative   0;  Mismatches   0;  Indels   0;  Gaps   0;

Qy          1 MDHLGASLWPQVGSLCLLLAGAAWAPPPNLPDPKFESKAALLAARGPEELLCFTERLEDL 60
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MDHLGASLWPQVGSLCLLLAGAAWAPPPNLPDPKFESKAALLAARGPEELLCFTERLEDL 60

Qy         61 VCFWEEAASAGVGPGNYSFSYQLEDEPWKLCRLHQAPTARGAVRFWCSLPTADTSSFVPL 120
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 VCFWEEAASAGVGPGNYSFSYQLEDEPWKLCRLHQAPTARGAVRFWCSLPTADTSSFVPL 120

Qy        121 ELRVTAASGAPRYHRVIHINEVVLLDAPVGLVARLADESGHVVLRWLPPPETPMTSHIRY 180
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 ELRVTAASGAPRYHRVIHINEVVLLDAPVGLVARLADESGHVVLRWLPPPETPMTSHIRY 180

Qy        181 EVDVSAGNGAGSVQRVEILEGRTECVLSNLRGRTRYTFAVRARMAEPSFGGFWSAWSEPV 240
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 EVDVSAGNGAGSVQRVEILEGRTECVLSNLRGRTRYTFAVRARMAEPSFGGFWSAWSEPV 240

Qy        241 SLLTPSDLDPLILTLSLILVVILVLLTVLALLSHRRALKQKIWPGIPSPESEFFEGLFTTH 300
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 SLLTPSDLDPLILTLSLILVVILVLLTVLALLSHRRALKQKIWPGIPSESEFFEGLFTTH 300

Qy        301 KGNFQLWLYQNDGCLWWSPCTPFTEDPPASLEVLSERCWGTMQAVEPGTDDEGPLLEPVG 360
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 KGNFQLWLYQNDGCLWWSPCTPFTEDPPASLEVLSERCWGTMQAVEPGTDDEGPLLEPVG 360

Qy        361 SEHAQDTYLVLDKWLLPRNPPSEDLPGPGGSVDIVAMDEGSEASSCSSALASKPSPEGAS 420
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 SEHAQDTYLVLDKWLLPRNPPSEDLPGPGGSVDIVAMDEGSEASSCSSALASKPSPEGAS 420

Qy        421 AASFEYTILDPSSQLLRPWTLCPELPPTPPHLKYLYLVVSDGISTDYSSGDSQGAQGGL 480
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 AASFEYTILDPSSQLLRPWTLCPELPPTPPHLKYLYLVVSDGISTDYSSGDSQGAQGGL 480

Qy        481 SDGPYSNPYENSLIPAAEPLPPSYVACS 508
              ||||||||||||||||||||||||||||
Db        481 SDGPYSNPYENSLIPAAEPLPPSYVACS 508

RESULT 4
AD005724
ID   AD005724 standard; protein; 508 AA.
XX
AC   AD005724;
XX
DT   15-JUL-2004 (first entry)
XX
DE   Human erythropoietin receptor (EPOR).
XX
KW   T cell; antiallergic; immunosuppressive; virucide; antibacterial;
KW   antiparasitic; cytostatic; gene therapy; human; receptor;
KW   erythropoietin receptor.
XX
OS   Homo sapiens.
XX
PN   WO2004032867-A2.
XX
PD   22-APR-2004.
XX
PF   09-OCT-2003; 2003WO-US032065.
XX
PR   09-OCT-2002; 2002US-0417102P.
PR   09-OCT-2002; 2002US-0417103P.
PR   09-OCT-2002; 2002US-0417243P.
PR   18-OCT-2002; 2002US-0419575P.
PR   08-NOV-2002; 2002US-0424777P.
PR   08-NOV-2002; 2002US-0424881P.
XX
PA   (TOLE-) TOLERRX INC.
XX
PI   Rao P, Szymanska G;
XX
DR   WPI; 2004-340801/31.
DR   N-PSDB; AD005723.
DR   GENBANK; 4557561.
XX
PT   Treating a condition that benefits from modulating regulatory or effector
PT   T cell function comprises administering an agent that modulates the
PT   expression or activity of a gene or polypeptide (e.g. PTGER2, TGFbetal,
PT   Jagged-1 or GPR-32).
XX
PS   Example; SEQ ID NO 24; 161pp; English.
XX
CC   The invention relates to treating a subject having a condition that
CC   benefits from modulating the balance of regulatory T cell function
CC   relative to effector T cell function, or vice versa, in a subject. The
CC   method involves administering an agent that modulates the expression or
CC   activity of a molecule selected from PTGER2 and TGFbetal, or Jagged-1,
CC   GPR-32, CD83, CD84, CD89, serotonin R, BY55, serotonin R2C, GPR63,
CC   histamine R-H4, GPR58, EPO-R, PSG-1, PSG-3, PSG-6, PSG-9, PDE-4d, and PI-
CC   3-related kinase to the subject, such that treatment occurs. The methods
CC   are useful for diagnosing, preventing or treating conditions
CC   characterized by a too-vigorous or weak effector T cell or regulatory T
```

CC   cell response to antigens associated with the condition, such as in an
CC   allergic response, an autoimmune disorder, a viral infection, a microbial
CC   infection, a parasitic infection or a tumour. The present sequence
CC   represents a human erythropoietin receptor (EPOR), preferentially
CC   expressed in regulatory T cells.
XX
SQ   Sequence 508 AA;

Query Match   100.0%; Score 2715; DB 8; Length 508;
Best Local Similarity   100.0%; Pred. No. 1.9e-216;
Matches   508; Conservative   0; Mismatches   0; Indels   0; Gaps   0;

Qy   1 MDHLGASLWPQVGSLCLLLAGAAWAPPPNLPDDPKFESKAALLAARGPEELLCFTERLEDL 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1 MDHLGASLWPQVGSLCLLLAGAAWAPPPNLPDDPKFESKAALLAARGPEELLCFTERLEDL 60

Qy   61 VCFWEEAASAGVGPGNYSFSYQLEDEPWKLCRLHQAPTARGAVRFWCSLPTADTSSFVPL 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61 VCFWEEAASAGVGPGNYSFSYQLEDEPWKLCRLHQAPTARGAVRFWCSLPTADTSSFVPL 120

Qy   121 ELRVTAASGAPRYHRVIHINEVVLLDAPVGLVARLADESGHVVLRWLPPPETPMTSHIRY 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   121 ELRVTAASGAPRYHRVIHINEVVLLDAPVGLVARLADESGHVVLRWLPPPETPMTSHIRY 180

Qy   181 EVDVSAGNGAGSVQRVEILEGRTECVLSNLRGRTRYTFAVRARMAEPSFGGFWSAWSEPV 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   181 EVDVSAGNGAGSVQRVEILEGRTECVLSNLRGRTRYTFAVRARMAEPSFGGFWSAWSEPV 240

Qy   241 SLLTPSDLDPLILTLSLILVVILVILVLLTVLALLSHRRALKQKIWPGIPSPESEFEGLFTTH 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   241 SLLTPSDLDPLILTLSLILVVILVILVLLTVLALLSHRRALKQKIWPGIPSPESEFEGLFTTH 300

Qy   301 KGNFQLWLYQNDGCLWWSPCTPFTEDPPASLEVLSERCWGTMQAVEPGTDDEGPLLEPVG 360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   301 KGNFQLWLYQNDGCLWWSPCTPFTEDPPASLEVLSERCWGTMQAVEPGTDDEGPLLEPVG 360

Qy   361 SEHAQDTYLVLDKWLLPRNPPSEDLPGPGGSVDIVAMDEGSEASSCSSALASKPSPEGAS 420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   361 SEHAQDTYLVLDKWLLPRNPPSEDLPGPGGSVDIVAMDEGSEASSCSSALASKPSPEGAS 420

Qy   421 AASFEYTILDPSSQLLRPWTLCPELPPTPHLKYLYLVVSDSGISTDYSSGDSQGQGGL 480
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   421 AASFEYTILDPSSQLLRPWTLCPELPPTPHLKYLYLVVSDSGISTDYSSGDSQGQGGL 480

Qy   481 SDGPYSNPYENSLIPAAEPLPPSYVACS 508
          ||||||||||||||||||||||||||||
Db   481 SDGPYSNPYENSLIPAAEPLPPSYVACS 508

RESULT 5
ADP12417
ID   ADP12417 standard; protein; 508 AA.
XX
AC   ADP12417;
XX
DT   12-AUG-2004 (first entry)
XX
DE   Protein encoded by mRNA of the invention #27.
XX
KW   transplant rejection; immune system; rheumatoid arthritis; lupus;
KW   inflammatory bowel disease; multiple sclerosis; HIV; AIDS.
XX
OS   Homo sapiens.
XX
PN   WO200404042346-A2.
XX
PD   21-MAY-2004.
XX
PF   24-APR-2003; 2003WO-US012946.
XX
PR   24-APR-2002; 2002US-00131831.
PR   20-DEC-2002; 2002US-00325899.
XX

PA   (EXPR-) EXPRESSION DIAGNOSTICS INC.
XX
PI   Wohlgemuth J, Fry K, Woodward R, Ly N, Prentice J, Morris M;
PI   Rosenberg S;
XX
DR   WPI; 2004-400724/37.
XX
PT   Diagnosing or monitoring transplant rejection, e.g. heart, kidney, liver,
PT   pancreas, pancreatic islet, lung, bone marrow or stem cell transplant
PT   rejection, in an individual, comprises detecting the expression level of
PT   the genes.
XX
PS   Claim 65; SEQ ID NO 2426; 1762pp; English.
XX
CC   The present invention relates to diagnosing or monitoring transplant
CC   rejection, e.g. cardiac or kidney transplant rejection, in an individual
CC   comprises detecting the expression level of one or more genes. The
CC   methods, system and kits are useful in diagnosing or monitoring
CC   transplant rejection, e.g. heart, kidney, liver, pancreas, pancreatic
CC   islet, lung, bone marrow or stem cell transplant rejection,
CC   xenotransplant rejection or mechanical organ replacement rejection, in an
CC   individual. The method is also useful in assessing the immune status of
CC   an individual. The methods are also useful in diagnosing and monitoring
CC   diseases that involve the immune system, e.g. rheumatoid arthritis,
CC   lupus, inflammatory bowel diseases, multiple sclerosis, HIV/AIDS or
CC   viral, bacterial or fungal infection. The present sequence represents a
CC   protein that is encoded by the mRNA of the invention.
XX
SQ   Sequence 508 AA;

Query Match   100.0%; Score 2715; DB 8; Length 508;
Best Local Similarity   100.0%; Pred. No. 1.9e-216;
Matches   508; Conservative   0; Mismatches   0; Indels   0; Gaps   0;

Qy   1 MDHLGASLWPQVGSLCLLLAGAAWAPPPNLPDDPKFESKAALLAARGPEELLCFTERLEDL 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1 MDHLGASLWPQVGSLCLLLAGAAWAPPPNLPDDPKFESKAALLAARGPEELLCFTERLEDL 60

Qy   61 VCFWEEAASAGVGPGNYSFSYQLEDEPWKLCRLHQAPTARGAVRFWCSLPTADTSSFVPL 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61 VCFWEEAASAGVGPGNYSFSYQLEDEPWKLCRLHQAPTARGAVRFWCSLPTADTSSFVPL 120

Qy   121 ELRVTAASGAPRYHRVIHINEVVLLDAPVGLVARLADESGHVVLRWLPPPETPMTSHIRY 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   121 ELRVTAASGAPRYHRVIHINEVVLLDAPVGLVARLADESGHVVLRWLPPPETPMTSHIRY 180

Qy   181 EVDVSAGNGAGSVQRVEILEGRTECVLSNLRGRTRYTFAVRARMAEPSFGGFWSAWSEPV 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   181 EVDVSAGNGAGSVQRVEILEGRTECVLSNLRGRTRYTFAVRARMAEPSFGGFWSAWSEPV 240

Qy   241 SLLTPSDLDPLILTLSLILVVILVILVLLTVLALLSHRRALKQKIWPGIPSPESEFEGLFTTH 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   241 SLLTPSDLDPLILTLSLILVVILVILVLLTVLALLSHRRALKQKIWPGIPSPESEFEGLFTTH 300

Qy   301 KGNFQLWLYQNDGCLWWSPCTPFTEDPPASLEVLSERCWGTMQAVEPGTDDEGPLLEPVG 360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   301 KGNFQLWLYQNDGCLWWSPCTPFTEDPPASLEVLSERCWGTMQAVEPGTDDEGPLLEPVG 360

Qy   361 SEHAQDTYLVLDKWLLPRNPPSEDLPGPGGSVDIVAMDEGSEASSCSSALASKPSPEGAS 420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   361 SEHAQDTYLVLDKWLLPRNPPSEDLPGPGGSVDIVAMDEGSEASSCSSALASKPSPEGAS 420

Qy   421 AASFEYTILDPSSQLLRPWTLCPELPPTPHLKYLYLVVSDSGISTDYSSGDSQGQGGL 480
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   421 AASFEYTILDPSSQLLRPWTLCPELPPTPHLKYLYLVVSDSGISTDYSSGDSQGQGGL 480

Qy   481 SDGPYSNPYENSLIPAAEPLPPSYVACS 508
          ||||||||||||||||||||||||||||
Db   481 SDGPYSNPYENSLIPAAEPLPPSYVACS 508

RESULT 6
ABM82037

```
ID   ABM82037 standard; protein; 508 AA.
XX
AC   ABM82037;
XX
DT   18-NOV-2004 (first entry)
XX
DE   Tumour-associated antigenic target (TAT) polypeptide PRO34323, SEQ:5259.
XX
KW   Tumour-associated antigenic target; TAT; human; overexpression; cancer;
KW   tumour; diagnosis; cell proliferative disorder; breast cancer;
KW   colorectal cancer; lung cancer; ovarian cancer; liver cancer;
KW   central nervous system cancer; bladder cancer; pancreatic cancer;
KW   cervical cancer; melanoma; leukaemia; hybridisation probe;
KW   gene therapy; cytostatic.
XX
OS   Homo sapiens.
XX
PN   WO2004030615-A2.
XX
PD   15-APR-2004.
XX
PF   29-SEP-2003; 2003WO-US028547.
XX
PR   02-OCT-2002; 2002US-0414971P.
XX
PA   (GETH ) GENENTECH INC.
XX
PI   Wu TD, Zhang Z, Zhou Y;
XX
DR   WPI; 2004-347921/32.
DR   N-PSDB; ACN40465.
XX
PT   New tumor-associated antigenic target polypeptides and nucleic acids,
PT   useful in preparing a medicament for treating or detecting a
PT   proliferative disorder, e.g. breast, lung, colorectal, ovarian or
PT   prostate cancer or tumor.
XX
PS   Claim 12; SEQ ID NO 5259; 7273pp; English.
XX
CC   The invention relates to human tumour-associated antigenic target (TAT)
CC   polypeptides, and their related nucleic acids. The TAT polypeptides are,
CC   overexpressed in cancer tissues compared to normal tissues, and may thus
CC   serve as effective targets for the diagnosis and treatment of cancer in
CC   mammals. The invention also relates to nucleic acid and polypeptide
CC   sequences at least 80% identical to the TAT nucleic acids and
CC   polypeptides; expression vectors and host cells comprising a TAT nucleic
CC   acid; an antibody specific for a TAT polypeptide; a peptide or organic
CC   molecule which binds to a TAT polypeptide; fusion proteins comprising a
CC   TAT polypeptide; and methods and compositions for the treatment or
CC   diagnosis of cancer in mammals. TAT polypeptides, nucleic acids,
CC   antibodies, antagonists, binding molecules and compositions are useful
CC   for diagnosing or treating a cell proliferative disorder associated with
CC   increased TAT expression, particularly cancers such as breast cancer,
CC   colorectal cancer, lung cancer, ovarian cancer, liver cancer, bladder
CC   cancer, pancreatic cancer, cervical cancer, cancers of the central
CC   nervous system, melanoma and leukaemia. TAT nucleic acids may further be
CC   used as hybridisation probes, in chromosome and gene mapping, in
CC   chromosome identification and in gene therapy. The present sequence
CC   represents a TAT polypeptide of the invention
XX
SQ   Sequence 508 AA;

  Query Match             100.0%;  Score 2715;  DB 8;  Length 508;
  Best Local Similarity   100.0%;  Pred. No. 1.9e-216;
  Matches  508;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 MDHLGASLWPQVGSLCLLLAGAAWAPPNLPDPKFESKAALLAARGPEELLCFTERLEDL 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MDHLGASLWPQVGSLCLLLAGAAWAPPNLPDPKFESKAALLAARGPEELLCFTERLEDL 60

Qy         61 VCFWEEAASAGVGPGNYSFSYQLEDEPWKLCRLHQAPTARGAVRFWCSLPTADTSSFVPL 120
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 VCFWEEAASAGVGPGNYSFSYQLEDEPWKLCRLHQAPTARGAVRFWCSLPTADTSSFVPL 120

Qy        121 ELRVTAASGAPRYRHVIHINEVVLLDAPVGLVARLADESGHVVLRWLPPPETPMTSHIRY 180
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 ELRVTAASGAPRYRHVIHINEVVLLDAPVGLVARLADESGHVVLRWLPPPETPMTSHIRY 180

Qy        181 EVDVSAGNGAGSVQRVEILEGRTECVLSNLRGRTRYTFAVRARMAEPSFGGFWSAWSEPV 240
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 EVDVSAGNGAGSVQRVEILEGRTECVLSNLRGRTRYTFAVRARMAEPSFGGFWSAWSEPV 240

Qy        241 SLLTPSDLDPLILTLSLILVVILVLLTVLALLSHRRALKQKIWPGIPSPESEFEGLFTTH 300
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 SLLTPSDLDPLILTLSLILVVILVLLTVLALLSHRRALKQKIWPGIPSPESEFEGLFTTH 300

Qy        301 KGNFQLWLYQNDGCLWWSPCTPFTEDPPASLEVLSERCWGTMQAVEPGTDDEGPLLEPVG 360
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 KGNFQLWLYQNDGCLWWSPCTPFTEDPPASLEVLSERCWGTMQAVEPGTDDEGPLLEPVG 360

Qy        361 SEHAQDTYLVLDKWLLPRNPPSEDLPGPGGSVDIVAMDEGSEASSCCSSALASKPSPEGAS 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 SEHAQDTYLVLDKWLLPRNPPSEDLPGPGGSVDIVAMDEGSEASSCCSSALASKPSPEGAS 420

Qy        421 AASFEYTILDPSSQLLRPWTLCPELPTTPPHLKYLYLVVSDSGISTDYSSGSDSQGAQGGL 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 AASFEYTILDPSSQLLRPWTLCPELPTTPPHLKYLYLVVSDSGISTDYSSGSDSQGAQGGL 480

Qy        481 SDGPYSNPYENSLIPAAEPLPPSYVACS 508
              ||||||||||||||||||||||||||||
Db        481 SDGPYSNPYENSLIPAAEPLPPSYVACS 508

RESULT 7
ABB09173
ID   ABB09173 standard; protein; 508 AA.
XX
AC   ABB09173;
XX
DT   01-JUL-2002 (first entry)
XX
DE   Human erythropoietin receptor SEQ ID NO:5.
XX
KW   Human; erythropoietin receptor; EPO receptor; fusion protein;
KW   erythropoietin receptor extracellular domain; signal transduction;
KW   receptor.
XX
OS   Homo sapiens.
XX
PN   US2002031806-A1.
XX
PD   14-MAR-2002.
XX
PF   30-JAN-1998;   98US-00016159.
XX
PR   16-JUN-1997;   97US-00876227.
XX
PA   (LEEJ/) LEE J Y.
XX
PI   Lee JY;
XX
DR   WPI; 2002-361181/39.
DR   N-PSDB; ABL51519.
XX
PT   New purified human erythropoietin receptor polypeptide, useful for
PT   studying ligand binding to erythropoietin receptor, as well as for
PT   analyzing and understanding receptor structure and signal transduction.
XX
PS   Example 1; Page 11-12; 19pp; English.
XX
CC   The present invention describes a purified human erythropoietin (EPO)
CC   receptor polypeptide, which consists essentially of amino acids 25-250 of
CC   the full length human EPO receptor protein. The human EPO receptor
CC   polypeptide is capable of binding human erythropoietin. Also described is
CC   a purified fusion protein comprising essentially of: (a) a first
```

```
CC    polypeptide segment having an amino terminus and a carboxyl terminus, the
CC    segment having a thrombin proteolytic cleavage site at the carboxyl
CC    terminus; and (b) a second polypeptide segment consisting essentially of
CC    the purified human EPO receptor, where the second polypeptide segment is
CC    covalently coupled to the carboxyl terminus of the first polypeptide
CC    segment. The human EPO receptor polypeptide is useful for studying ligand
CC    binding to the EPO receptor and for quantitating the amounts of the EPO
CC    receptor. The polypeptide is also useful for analysing and understanding
CC    receptor structure and signal transduction. The present sequence
CC    represents human EPO receptor from the present invention
XX
SQ    Sequence 508 AA;

Query Match        99.7%;   Score 2708;  DB 5;   Length 508;
Best Local Similarity  99.8%;   Pred. No. 7.1e-216;
Matches   507;  Conservative   0;  Mismatches   1;  Indels   1;  Gaps   0;

Qy      1  MDHLGASLWPQVGSLCLLLAGAAWAPPNLPDPKFESKAALLAARGPEELLCFTERLEDL   60
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1  MDHLGASLWPQVGSLCLLLAGAAWAPPNLPDPKFESKAALLAARGPEELLCFTERLEDL   60

Qy     61  VCFWEEAASAGVGPGNYSFSYQLEDEPWKLCRLHQAPTARGAVRFWCSLPTADTSSFVPL  120
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61  VCFWEEAASAGVGPGNYSFSYQLEDEPWKLCRLHQAPTARGAVRFWCSLPTADTSSFVPL  120

Qy    121  ELRVTAASGAPRYHRVHINEVVLLDAPVGLVARLADESGHVVLRWLPPPETPMTSHIRY  180
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121  ELRVTAASGAPRYHRVHINEVVLLDAPVGLVARLADESGHVVLRWLPPPETPMTSHIRY  180

Qy    181  EVDVSAGNGAGSVQRVEILEGRTECVLSNLRGRTRYTFAVRARMAEPSFGGFWSAWSEPV  240
           ||||||||.||||||||||||||||||||||||||||||||.|||||||||||||||||
Db    181  EVDVSAGNGAGSVQRVEILEGRTECVLSNLRGRTRYTFAVLARMAEPSFGGFWSAWSBPV  240

Qy    241  SLLTPSDLDPLILTLSLILVVLILVVILTVLALLSHRRALKQKIWPGIPSPESEFEGLFTTH  300
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    241  SLLTPSDLDPLILTLSLILVVILVVILTVLALLSHRRALKQKIWPGIPSPESEFEGLFTTH  300

Qy    301  KGNFQLWLYQNDGCLWWSPCTPFTEDDPPASLEVLSERCWGTMQAVEPGTDDEGPLLEPVG  360
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    301  KGNFQLWLYQNDGCLWWSPCTPFTEDDPPASLEVLSERCWGTMQAVEPGTDDEGPLLEPVG  360

Qy    361  SEHAQDTYLVLDKWLLPRNPPSEDLPGPGGSVDIVAMDEGSEASSCCSSALASKPSPEGAS  420
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    361  SEHAQDTYLVLDKWLLPRNPPSEDLPGPGGSVDIVAMDEGSEASSCCSSALASKPSPEGAS  420

Qy    421  AASFEYTILDPSSQLLRPWTLCPELPPTPPHLKYLYLVVSDSGISTDYSSSGDSQGAQGGL  480
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    421  AASFEYTILDPSSQLLRPWTLCPELPPTPPHLKYLYLVVSDSGISTDYSSSGDSQGAQGGL  480

Qy    481  SDGPYSNPYENSLIPAAEPLPPSYVACS  508
           ||||||||||||||||||||||||||||
Db    481  SDGPYSNPYENSLIPAAEPLPPSYVACS  508

RESULT 8
ADM93422
ID    ADM93422 standard; protein; 508 AA.
XX
AC    ADM93422;
XX
DT    01-JUL-2004  (first entry)
XX
DE    Human NOVX polypeptide #27.
XX
KW    Human; NOVX; congenital heart defect; cardiomyopathy; atherosclerosis;
KW    hypertension; pulmonary stenosis; scleroderma; adenocarcinoma;
KW    haemophilia; graft-versus-host disease; cancer;
KW    neurodegenerative disorder; Alzheimer's disease; Parkinson's disease;
KW    multiple sclerosis; diabetes; obesity; bronchial asthma;
KW    acquired immunodeficiency syndrome; AIDS; Crohn's disease;
KW    infectious disease; anorexia; immune disorder.
XX
OS    Homo sapiens.
```

```
XX
PN    US2004067882-A1.
XX
PD    08-APR-2004.
XX
PF    05-NOV-2002; 2002US-00287971.
XX
PR    22-OCT-2001; 2001US-00035568.
PR    05-NOV-2001; 2001US-0338626P.
PR    06-NOV-2001; 2001US-0333072P.
PR    09-NOV-2001; 2001US-0345398P.
PR    09-NOV-2001; 2001US-0348283P.
PR    15-NOV-2001; 2001US-0335610P.
PR    21-NOV-2001; 2001US-0332152P.
PR    28-NOV-2001; 2001US-0333912P.
PR    29-NOV-2001; 2001US-00997425.
PR    29-NOV-2001; 2001US-0334300P.
PR    04-DEC-2001; 2001US-0336576P.
PR    05-FEB-2002; 2002US-0354807P.
PR    15-MAY-2002; 2002US-0380968P.
PR    16-MAY-2002; 2002US-0381043P.
PR    02-JUL-2002; 2002US-0393148P.
PR    02-JUL-2002; 2002US-0393262P.
PR    06-AUG-2002; 2002US-0401479P.
PR    06-AUG-2002; 2002US-0401626P.
PR    07-AUG-2002; 2002US-0401593P.
PR    07-AUG-2002; 2002US-0401695P.
PR    26-AUG-2002; 2002US-0406181P.
XX
PA    (ALSO/)  ALSOBROOK J P.
PA    (ALVA/)  ALVAREZ E.
PA    (ANDE/)  ANDERSON D W.
PA    (BARO/)  BARON M.
PA    (BOLD/)  BOLDOG F L.
PA    (BURG/)  BURGESS C E.
PA    (CASM/)  CASMAN S J.
PA    (CHAP/)  CHAPOVAL A.
PA    (DHAN/)  DHANABAL M.
PA    (EDIN/)  EDINGER S R.
PA    (EISE/)  EISEN A.
PA    (ELLE/)  ELLERMAN K.
PA    (ETTE/)  ETTENBERG S.
PA    (GANG/)  GANGOLLI E A.
PA    (GERL/)  GERLACH V.
PA    (GORM/)  GORMAN L.
PA    (GROS/)  GROSSE W M.
PA    (GUOX/)  GUO X.
PA    (HACK/)  HACKETT C.
PA    (JIWW/)  JI W.
PA    (KEKU/)  KEKUDA R.
PA    (KHRA/)  KHRAMTSOV N V.
PA    (LEPL/)  LEPLEY D M.
PA    (LILL/)  LI L.
PA    (MACD/)  MACDOUGALL J R.
PA    (MALY/)  MALYANKAR U M.
PA    (MAZU/)  MAZUR A.
PA    (MCQU/)  MCQUEENEY K.
PA    (MEZE/)  MEZES P S.
PA    (MILL/)  MILLER C E.
PA    (MILL/)  MILLET I.
PA    (MISH/)  MISHRA V.
PA    (PADI/)  PADIGARU M.
PA    (PATT/)  PATTURAJAN M.
PA    (PENA/)  PENA C E A.
PA    (PEYM/)  PEYMAN J A.
PA    (RAST/)  RASTELLI L.
PA    (RIEG/)  RIEGER D K.
PA    (ROTH/)  ROTHENBERG M E.
PA    (SHEN/)  SHENOY S G.
PA    (SHIM/)  SHIMKETS R A.
PA    (SMIT/)  SMITHSON G.
PA    (SPAD/)  SPADERNA S K.
PA    (STAR/)  STARLING G.
```

```
PA  (SPYT/)  SPYTEK K A.
PA  (STON/)  STONE D J.
PA  (TCHE/)  TCHERNEV V T.
PA  (TWOM/)  TWOMLOW N.
PA  (VERN/)  VERNET C A M.
PA  (ZERH/)  ZERHUSEN B D.
PA  (VOSS/)  VOSS E Z.
PA  (ZHON/)  ZHONG M.
XX
PI  Alsobrook JP,  Alvarez E,  Anderson DW,  Baron M,  Boldog FL;
PI  Burgess CE,  Casman SJ,  Chapoval A,  Dhanabal M,  Edinger SR,     Eisen A;
PI  Ellerman K,  Ettenberg S,  Gangolli EA,  Gerlach V,  Gorman L;
PI  Grosse WM,  Guo X,  Hackett C,  Ji W,  Kekuda R,     Khramtsov NV;
PI  Lepley DM,  Li L,  Macdougall JR,  Malyankar UM,  Mazur A,  Mcqueeney K;
PI  Mezes PS,  Miller CE,  Millet I,  Mishra V,  Padigaru M,  Patturajan M;
PI  Pena CEA,  Peyman JA,  Rastelli L,  Rieger DK,  Rothenberg ME;
PI  Shenoy SG,  Shinkets RA,  Smithson G,  Spaderna SK,  Starling G;
PI  Spytek KA,  Stone DJ,  Tchernev VT,  Twomlow N,  Vernet CAM;
PI  Zerhusen BD,  Voss EZ,  Zhong M;
XX
DR  WPI; 2004-355303/33.
DR  N-PSDB; ADM93421.
XX
PT  Novel isolated NOVX polypeptide useful treating or preventing disorders
PT  or syndromes such as Alzheimer's disease, Parkinson's disease, multiple
PT  sclerosis, diabetes, obesity, cancer, bronchial asthma, Crohn's disease.
XX
PS  Claim 2; SEQ ID NO 54; 330pp; English.
XX
CC  The invention relates to human NOVX polypeptides and the polynucleotides
CC  encoding them. The NOVX polypeptides and polynucleotides are useful for
CC  determining the presence of or predisposition to a disease associated
CC  with altered levels of the sequences in a mammalian subject, and for
CC  treating or preventing a pathology associated with NOVX. The
CC  polypeptides, polynucleotides and antibodies that bind immunospecifically
CC  to the polypeptides are useful for treating or preventing disorders or
CC  syndromes such as congenital heart defects, cardiomyopathy,
CC  atherosclerosis, hypertension, pulmonary stenosis, scleroderma,
CC  adenocarcinoma, haemophilia, graft-versus-host disease, cancer,
CC  neurodegenerative disorders, Alzheimer's disease, Parkinson's disease,
CC  multiple sclerosis, diabetes, obesity, bronchial asthma, acquired
CC  immunodeficiency syndrome (AIDS), Crohn's disease, infectious disease,
CC  anorexia and immune disorders. This sequence represents a human NOVX
CC  polypeptide of the invention. Note: The sequence data for this patent is
CC  also available from USPTO at seqdata.uspto.gov/sequence.html.
XX
SQ  Sequence 508 AA;
XX
Query Match             99.2%;  Score 2693;  DB 8;  Length 508;
Best Local Similarity   99.2%;  Pred. No. 1.3e-214;
Matches  504;  Conservative    1;  Mismatches    3;  Indels    0;  Gaps    0;

Qy        1 MDHLGASLWPQVGSLCLLLAGAAWAPPPNLPDPKFESKAALLAARGPEELLCFTERLEDL 60
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1 MDHLGASLWPQVGSLCLLPAGAAWAPPPNLPDPKFESKAALLAARGPEELLCFTERLEDL 60

Qy       61 VCFWEEAASAGVGPGNYSFSYQLEDEPWKLCRLHQAPTARGAVRFWCSLPTADTSSFVPL 120
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       61 VCFWEEAASAGVGPGNYSFSYQLEDEPWKLCRLHQAPTARGAVRFWCSLPTADTSSFVPL 120

Qy      121 ELRVTAASGAPRYRHRVIHINEVVLLDAPVGLVARLADESGHVVLRWLPPETPMTSHIRY 180
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      121 ELRVTAASGAPRYRHRVIHINEVVLLDAPVGLVARLADESGHVVLRWLPPETPMTSHIRY 180

Qy      181 EVDVSAGNGAGSVQRVEILEGRTECVLSNLRGRTRYTFAVRARMAEPSFGGFWSAWSEPV 240
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      181 AVDVSAGNGAGSVQRVKILEGRTECVLSNLRGRTRYTFAVRARMAEPSFGGFWSAWSEPV 240

Qy      241 SLLTPSDLDPLILTLSLIIVVILVLLLTVLALLSHRRALKQKIWPGIPSPESEFEGLFTTH 300
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      241 SLLTPSDLDPLILTLSLIIVVILVLLLTVLALLSHRRALKQKIWPGIPSPESEFEGLFTTH 300

Qy      301 KGNFQLWLYQNDGCLWWSPCTPFTEDPPASLEVLSERCWGTMQAVEPGTDDEGPLLEPVG 360
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      301 KGNFQLWLYQNDGCLWWSPCTPFTEDPPASLEVLSERCWGTMQAVEPGTDDEGPLLEPVG 360

Qy      361 SEHAQDTYLVLDKWLLPRNPPSEDLPGPGGSVDIVAMDEGSEASSCSSALASKPSPEGAS 420
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      361 SEHAQDTYLVLDKWLLPRNPPSEDLPGPGGSVDIVAMDEGSEASSCSSALASKPSPEGAS 420

Qy      421 AASFEYTILDPSSQLLRPWTLCPELPPTPPHLKYLYLVVSDSGISTDYSSGDSGQAQGGL 480
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      421 AASFEYTILDPSQLLRPWTLCPELPPTPPHLKYLYLVVSDSGISTDYSSGDSGQAQGGL 480

Qy      481 SDGPYSNPYENSLIPAAEPLPPSYVACS 508
            |||||||||||||||||||||||||||
Db      481 SDGPYSNPYENSLIPAAEPLPPSYVACS 508


RESULT 9
ADE28677
ID   ADE28677 standard; protein; 508 AA.
XX
AC   ADE28677;
XX
DT   29-JAN-2004  (first entry)
XX
DE   Human NOV15d protein - SEQ ID 54.
XX
KW   NOVX; antidiabetic; anorectic; cardiant; hypotensive;
KW   antiarteriosclerotic; virucide; antibacterial; fungicide; protozoacide;
KW   nootropic; neuroprotective; antiparkinsonian; anticonvulsant;
KW   osteopathic; antiarthritic; antiinflammatory; dermatological;
KW   antiasthmatic; antiallergic; metabolic; diabetes; obesity; infectious;
KW   anorexia; cancer; cardiovascular; hypertension; atherosclerosis;
KW   neurodegenerative; Alzheimer's disease; Parkinson's; epilepsy; immune;
KW   osteoarthritis; haemopoietic; inflammatory skin; asthma; dyslipidaemia;
KW   neurogenesis; cell differentiation; proliferation; haemopoiesis;
KW   wound healing; angiogenesis; gene therapy; chromosome mapping;
KW   tissue typing; human; NOV.
XX
OS   Homo sapiens.
XX
PN   WO200304030-A2.
XX
PD   15-MAY-2003.
XX
PF   05-NOV-2002; 2002WO-US035536.
XX
PR   05-NOV-2001;  2001US-0338626P.
PR   05-DEC-2001;  2001US-0336600P.
PR   07-DEC-2001;  2001US-0338285P.
PR   12-DEC-2001;  2001US-0341346P.
PR   17-DEC-2001;  2001US-0341477P.
PR   17-DEC-2001;  2001US-0341540P.
PR   20-DEC-2001;  2001US-0342592P.
PR   27-DEC-2001;  2001US-0342297P.
PR   31-DEC-2001;  2001US-0344903P.
PR   17-APR-2002;  2002US-0373288P.
PR   15-MAY-2002;  2002US-0380981P.
PR   17-MAY-2002;  2002US-0381495P.
PR   28-MAY-2002;  2002US-0383534P.
PR   28-MAY-2002;  2002US-0383744P.
PR   29-MAY-2002;  2002US-0383829P.
PR   29-MAY-2002;  2002US-0384024P.
PR   07-AUG-2002;  2002US-0401788P.
PR   26-AUG-2002;  2002US-0406353P.
PR   31-OCT-2002;  2002US-00287971.
XX
PA   (CURA-) CURAGEN CORP.
XX
PI   Alsobrook JP,  Alvarez E,  Anderson DW,  Baron M,   Boldog FL;
PI   Burgess CE,  Casman SJ,  Chapoval A,  Dhanabal M,  Edinger SR,    Eisen A;
PI   Ellerman K,  Ettenberg S,  Gangolli EA,  Gerlach VL,  Gorman L;
PI   Grosse WM,  Guo X,  Hackett C,  Ji W, Kekuda R,   Khramtsov NV;
PI   Lepley DM,  Li L,  Macdougall JR,  Malyankar UM,  Mazur A,  Mcqueeney K;
```

PI   Mezes PS,   Miller CE,   Millet I,   Mishra VS,   Padigaru M,   Patturajan M;
PI   Pena CEA,   Peyman JA,   Rastelli L,   Rieger DK,   Shenoy SG,   Shimkets RA;
PI   Smithson G,   Starling G,   Spytek KA,   Stone DJ,   Tchernev VT,   Twomlow N;
PI   Vernet CAM,   Zerhusen BD,   Zhong M;
XX
DR   WPI; 2003-441555/41.
DR   N-PSDB; ADE28676.
XX
PT   New isolated NOVX polypeptides and polynucleotides, useful for
PT   preventing, diagnosing or treating NOVX-associated disorders, e.g.
PT   osteoarthritis, obesity, atherosclerosis, cancer, Parkinson's disease,
PT   asthma, or infections.
XX
PS   Claim 1; SEQ ID NO 54; 447pp; English.
XX
CC   The invention relates to a novel isolated NOVX polypeptide. The
CC   polypeptide of the invention demonstrates, antidiabetic, anorectic,
CC   cardiant, hypotensive, antiarteriosclerotic, virucide, antibacterial,
CC   fungicide, protozoacide, nootropic, neuroprotective, antiparkinsonian,
CC   anticonvulsant, osteopathic, antiarthritic, antiinflammatory,
CC   dermatological, antiasthmatic and antilipaemic activities. The
CC   polypeptides, nucleic acid molecules and antibodies may be useful for
CC   treating or diagnosing diseases including metabolic disorders such as
CC   diabetes and obesity, infectious diseases, anorexia, cancer,
CC   cardiovascular diseases including hypertension and atherosclerosis,
CC   neurodegenerative disorders such as Alzheimer's disease, Parkinson's
CC   disease and epilepsy, immune disorders e.g. osteoarthritis, haemopoietic
CC   disorders, inflammatory skin disorders, asthma and dyslipidaemia.
CC   Furthermore, the nucleic acids and polypeptides may also be used to
CC   identify molecules that modulate or inhibit neurogenesis, cell
CC   differentiation and proliferation, haemopoiesis, wound healing and
CC   angiogenesis, as well as in gene therapy. Finally, the nucleic acids may
CC   be used as hybridisation probes, in chromosome mapping, tissue typing,
CC   preventive medicine and pharmacogenomics. The current sequence is that of
CC   the human NOV protein of the invention.
XX
SQ   Sequence 508 AA;

Query Match   99.1%; Score 2691; DB 7; Length 508;
Best Local Similarity   99.0%; Pred. No. 1.8e-214;
Matches 503; Conservative   2; Mismatches   3; Indels   3; Gaps   0;

Qy   1 MDHLGASLWPQVGSLCLLLAGAAWAPPPNLPDPKFESKAALLAARGPEELLCFTERLEDL 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1 MDHLGASLWPQVGSLCLLLPAGAAWAPPPNLPDPKFESKAALLAARGPEELLCFTERLEDL 60

Qy   61 VCFWEEAASAGVGPGNYSFSYQLEDEPWKLCRLHQAPTARGAVRFWCSLPTADTSSFVPL 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61 VCFWEEAASAGVGPGNYSFSYQLEDEPWKLCRLHQAPTARGAVRFWCSLPTADTSSFVPL 120

Qy   121 ELRVTAASGAPRYRVIHINEVVLLDAPVGLVARLADESGHVVLRWLPPETPMTSHIRY 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   121 ELRVTAASGAPRYRVIHINEVVLLDAPVGLVARLADESGHVVLRWLPPETPMTSHIRY 180

Qy   181 EVDVSAGNGAGSVQRVEILEGRTECVLSNLRGRTRYTFAVRARMAEPSFGGFWSAWSEPV 240
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   181 AVDVSAGNGAGSVQRVKILEGRTECVLSNIRGRTRYTFAVRARMAEPSFGGFWSAWSEPV 240

Qy   241 SLLTPSDLDPLIITLSLILVVILVVLLTVLALLSHRRALKQKIWPGIPSPESEFEGLFTTH 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   241 SLLTPSDLDPLIITLSLILVVILVVILVLLTVLALLSHRRALKQKIWPGIPSPESEFEGLFTTH 300

Qy   301 KGNFQLWLYQNDGCLWWSPCTPFTEDPPASLEVLSERCWGTMQAVEPGTDDEGPLLEPVG 360
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   301 KGNFQLWLYQNDGCLWWSPCTPFTEDPPASLEVLSERCWGTMQAVEPGTDDEGPLLEPVG 360

Qy   361 SEHAQDTYLVLDKWLLPRNPPSEDLPGPGGSVDIVAMDEGSEASSCSSALASKPSPEGAS 420
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   361 SEHAQDTYLVLDKWLLPRNPPSEDLPGPGGSVDIVAMDEGSEASSCSSALASKPSPEGAS 420

Qy   421 AASFEYTILDPSSQLLRPWTLCPELPPTPPHLKYLYLVVSDSGISTDYSSGDSQGAQGGL 480
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   421 AASFEYTILDPSPQLLRPWTLCPELPPTPPHLKYLYLVVSDSGISTDYSSGDSQGAQGGL 480

Qy   481 SDGPYSNPYENSLIPAAEPLPPSYVACS 508
          |||||||||||||||||||||||||||
Db   481 SDGPYSNPYENSLIPAAEPLPPSYVACS 508

RESULT 10
AAR47518
ID   AAR47518 standard; protein; 508 AA.
XX
AC   AAR47518;
XX
DT   25-MAR-2003 (revised)
DT   24-JUN-1994 (first entry)
XX
DE   Human EPO receptor.
XX
KW   Erythropoietin receptor; recombinant; murine; anaemia.
XX
OS   Homo sapiens.
XX
FH   Key   Location/Qualifiers
FT   Peptide   1..24
FT   /note="signal peptide"
FT   Protein   25..508
FT   /note="mature EPO receptor"
FT   Region   251..272
FT   /note="putative transmembrane domain"
XX
PN   US5278065-A.
XX
PD   11-JAN-1994.
XX
PF   25-MAR-1991; 91US-00678877.
XX
PR   03-FEB-1989; 89US-00306503.
XX
PA   (GEMY) GENETICS INST INC.
PA   (CHIL-) CHILDRENS MEDICAL CENT.
PA   (WHED) WHITEHEAD INST BIOMEDICAL RES.
XX
PI   D'andrea A, Wong GG, Jones SS;
XX
DR   WPI; 1994-025409/03.
DR   N-PSDB; AAQ53995.
XX
PT   Recombinant DNA encoding erythropoietin receptor - used to develop prods.
PT   for study, treatment or diagnosis of disorders in which receptor is
PT   dysfunctional.
XX
PS   Disclosure; Fig 9; 24pp; English.
XX
CC   Mouse erythroleukaemia (MEL) cells were used to construct a cDNA library.
CC   The cDNA was used to transfect COS-1 cells and these were screened for
CC   radioiodinated erythropoietin (EPO) binding to isolate cDNA encoding the
CC   EPO receptor. This cDNA was used as a probe to screen a human genomic
CC   cDNA library to obtain DNA encoding the human EPO receptor. The cDNA may
CC   be used to study, treat or diagnose disorders in which the EPO receptor
CC   is dysfunctional.The EPO receptor may also be used to raise antibodied or
CC   for treating hypersensitivity to EPO or who have elevated levels of EPO.
CC   The prod. is pref. used for treating anaemias, primary proliferative
CC   polycythemia and secondary polycythemia. See also AAR47517. (Updated on
CC   25-MAR-2003 to correct PF field.)
XX
SQ   Sequence 508 AA;

Query Match   99.1%; Score 2690; DB 2; Length 508;
Best Local Similarity   98.8%; Pred. No. 2.2e-214;
Matches 502; Conservative   3; Mismatches   3; Indels   0; Gaps   0;

Qy   1 MDHLGASLWPQVGSLCLLLAGAAWAPPPNLPDPKFESKAALLAARGPEELLCFTERLEDL 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1 MDHLGASLWPQVGSLCLLLAGAAWAPPPNLPDPKFESKAALLAARGPEELLCFTERLEDL 60

```
Qy    61  VCFWEEAASAGVGPGNYSFSYQLEDEPWKLCRLHQAPTARGAVRFWCSLPTADTSSFVPL 120
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    61  VCFWEEGASAGVGPGNYSFSYQLEDEPWKLCRLHQAPTARGAVRFWCSLPTADTSSFVPL 120

Qy   121  ELRVTAASGAPRYHRVIHINEVVLLDAPVGLVARLADESGHVVLRWLPPPETPMTSHIRY 180
              |||||||||||||||||||||:|||||||||||||||||||||||||||||||||||
Db   121  ELRVTAASGAPRYHRVIHMNEVVLLDAPVGLVARLADESGHVVLRWLPPPETPMTSHIRY 180

Qy   181  EVDVSAGNGAGSVQRVEILEGRTECVLSNLRGRTRYTFAVRARMAEPSFGGFWSAWSEPV 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   181  EVDVSAGNGAGSVQRVEILEGRTECVLSNLRGRTRYTFAVRARMAEPSFGGFWSAWSEPV 240

Qy   241  SLLTPSDLDPLILTLSLILVVILVILLTVLALLSHRRALKQKIWPGIPSPESEFBGLFTH 300
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   241  SLLTPSDLDPLILTLSLILVVILVILLTVLALLSHRRALKQKIWPGIPSPESEFEGLFTH 300

Qy   301  KGNFQLWLYQNDGCLWWSPCTPFTEDPPASLEVLSERCWGTMQAVEPGTDDEGPLLEPVG 360
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   301  KGNFQLWLYQNDGCLWWSPCTPLTEDPPASLEVLSERCWGTMQAVEPGTDDEGPLLEPVG 360

Qy   361  SEHAQDTYLVLDKWLLPRNPPSEDLPGPGGSVDIVAMDEGSEASSCSSALASKPSPEGAS 420
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   361  SEHAQDTYLVLDKWLLPRNPPSEDLPGPGGSVEIVAMDEGSEASSCSSALASKPSPEGAS 420

Qy   421  AASFEYTILDPSSQLLRPWTLCPELPPTPPHLKYLYLVVSDSGISTDYSSGDSQGAQGGL 480
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   421  AASFEYTILDPSSQLLRPWTLCPELPPTPPHLKYLYLVVCDSGMSTDYSSGDSQGAQGGL 480

Qy   481  SDGPYSNPYENSLIPAAEPLPPSVVACS 508
              |||||||||||||||||||||||||||
Db   481  SDGPYSNPYENSLIPAAEPLPPSVVACS 508
```

RESULT 11
ADE28675
ID   ADE28675 standard; protein; 458 AA.
XX
AC   ADE28675;
XX
DT   29-JAN-2004 (first entry)
XX
DE   Human NOV15c protein - SEQ ID 52.
XX
KW   NOVX; antidiabetic; anorectic; cardiant; hypotensive;
KW   antiarteriosclerotic; virucide; antibacterial; fungicide; protozoacide;
KW   nootropic; neuroprotective; antiparkinsonian; anticonvulsant;
KW   osteopathic; antiarthritic; antiinflammatory; dermatological;
KW   antiasthmatic; antilipaemic; metabolic; diabetes; obesity; infectious;
KW   anorexia; cancer; cardiovascular; hypertension; atherosclerosis;
KW   neurodegenerative; Alzheimer's disease; Parkinson's; epilepsy; immune;
KW   osteoarthritis; haemopoietic; inflammatory skin; asthma; dyslipidaemia;
KW   neurogenesis; cell differentiation; proliferation; haemopoiesis;
KW   wound healing; angiogenesis; gene therapy; chromosome mapping;
KW   tissue typing; human; NOV.
XX
OS   Homo sapiens.
XX
PN   WO2003040330-A2.
XX
PD   15-MAY-2003.
XX
PF   05-NOV-2002; 2002WO-US035536.
XX
PR   05-NOV-2001; 2001US-0338626P.
PR   05-DEC-2001; 2001US-0336600P.
PR   07-DEC-2001; 2001US-0338285P.
PR   12-DEC-2001; 2001US-0341346P.
PR   17-DEC-2001; 2001US-0341477P.
PR   17-DEC-2001; 2001US-0341540P.
PR   20-DEC-2001; 2001US-0342592P.
PR   27-DEC-2001; 2001US-0344297P.
PR   31-DEC-2001; 2001US-0344903P.

PR   17-APR-2002; 2002US-0372288P.
PR   15-MAY-2002; 2002US-0380981P.
PR   17-MAY-2002; 2002US-0381495P.
PR   28-MAY-2002; 2002US-0383534P.
PR   28-MAY-2002; 2002US-0383744P.
PR   29-MAY-2002; 2002US-0383829P.
PR   29-MAY-2002; 2002US-0384024P.
PR   07-AUG-2002; 2002US-0401788P.
PR   26-AUG-2002; 2002US-0406353P.
PR   31-OCT-2002; 2002US-00287971
XX
PA   (CURA-) CURAGEN CORP.
XX
PI   Alsobrook JP, Alvarez E, Anderson DW, Baron M, Boldog FL;
PI   Burgess CE, Casman SJ, Chapoval A, Dhanabal M, Edinger SR, Eisen A;
PI   Ellerman K, Ettenberg S, Gangolli EA, Gerlach VL, Gorman L;
PI   Grosse WM, Guo X, Hackett C, Ji W, Kekuda R, Khramtsov NV;
PI   Lepley DM, Li L, Macdougall JR, Malyankar UM, Mazur A, Mcqueeney K;
PI   Mezes PS, Miller CE, Millet I, Mishra VS, Padigaru M, Patturajan M;
PI   Pena CEA, Peyman JA, Rastelli L, Rieger DK, Shenoy SG, Shimkets RA;
PI   Smithson G, Starling G, Spytek KA, Stone DJ, Tchernev VT, Twomlow N;
PI   Vernet CAM, Zerhusen BD, Zhong M;
XX
DR   WPI; 2003-441555/41.
DR   N-PSDB; ADE28674.
XX
PT   New isolated NOVX polypeptides and polynucleotides, useful for
PT   preventing, diagnosing or treating NOVX-associated disorders, e.g.
PT   osteoarthritis, obesity, atherosclerosis, cancer, Parkinson's disease,
PT   asthma, or infections.
XX
PS   Claim 1; SEQ ID NO 52; 447pp; English.
XX
CC   The invention relates to a novel isolated NOVX polypeptide. The
CC   polypeptide of the invention demonstrates, antidiabetic, anorectic,
CC   cardiant, hypotensive, antiarteriosclerotic, virucide, antibacterial,
CC   fungicide, protozoacide, nootropic, neuroprotective, antiparkinsonian,
CC   anticonvulsant, osteopathic, antiarthritic, antiinflammatory,
CC   dermatological, antiasthmatic and antilipaemic activities. The
CC   polypeptides, nucleic acid molecules and antibodies may be useful for
CC   treating or diagnosing diseases including metabolic disorders such as
CC   diabetes and obesity, infectious diseases, anorexia, cancer,
CC   cardiovascular diseases including hypertension and atherosclerosis,
CC   neurodegenerative disorders such as Alzheimer's disease, Parkinson's
CC   disease and epilepsy, immune disorders e.g. osteoarthritis, haemopoietic
CC   disorders, inflammatory skin disorders, asthma and dyslipidaemia.
CC   Furthermore, the nucleic acids and polypeptides may also be used to
CC   identify molecules that modulate or inhibit neurogenesis, cell
CC   differentiation and proliferation, haemopoiesis, wound healing and
CC   angiogenesis, as well as in gene therapy. Finally, the nucleic acids may
CC   be used as hybridisation probes, in chromosome mapping, tissue typing,
CC   preventive medicine and pharmacogenomics. The current sequence is that of
CC   the human NOV protein of the invention.
XX
SQ   Sequence 458 AA;

Query Match   88.7%; Score 2407; DB 7; Length 458;
Best Local Similarity   89.2%; Pred. No. 6.1e-191;
Matches   453; Conservative   1; Mismatches   1; Indels   4; Gaps   1;

```
Qy     1  MDHLGASLWPQVGSLCLLLAGAAWAPPPNLPDPKFESKAALLAARGPEELLCFTERLEDL 60
             |||||||||||||||||||||||||||| |||||||||||||||||||||||||||||:|
Db     1  MDHLGASLWPQVGSLCLLLAGAAWAPPPNLPDPKFESKAALLAARGPEELLCFTERLGDL 60

Qy    61  VCFWEEAASAGVGPGNYSFSYQLEDEPWKLCRLHQAPTARGAVRFWCSLPTADTSSFVPL 120
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    61  VCFWEEGASAGVGPGNYSFSYQLEDEPWKLCRLHQAPTARGAVRFWCSLPTADTSSFVPL 120

Qy   121  ELRVTAASGAPRYHRVIHINEVVLLDAPVGLVARLADESGHVVLRWLPPPETPMTSHIRY 180
             ||||||||||||||||||:|||||||||||||||||||||||||||||||||||||||||
Db   121  ELRVTAASGAPRYHRVIHMNEVVLLDAPVGLVARLADESGHVVLRWLPPPETPMTSHIRY 180

Qy   181  EVDVSAGNGAGSVQRVEILEGRTECVLSNLRGRTRYTFAVRARMAEPSFGGFWSAWSEPV 240
```

```
Db        181 |||||||||   ||||||||||||||||   |||||||||||||
              EVDVSAGNGAGSVQRVEILEGRTECVLSNLRGRTRYTFAVRTRMAEPSFGGFWSAWSEPV 240

Qy        241 SLLTPSDLDPLILTLSLILVVILVLLTVLALLSHRRALKQKIWPGIPSPESEFEGLFTTH 300
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 SLLTPSDLDPLILTLSLILVVILVLLTVLALLSHRRALKQKIWPGIPSPESEFEGLFTTH 300

Qy        301 KGNFQLWLYQNDGCLWWSPCTPFTEDPPASLEVLSERCWGTMQAVEPGTDDEGPLLEPVG 360
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 KGNFQLWLYQNDGCLWWSPCTPFTEDPPASLEVLSERCWGTMQAVEPGTDDEGPLLEPVG 360

Qy        361 SEHAQDTYLVLDKWLLPRNPPSEDLPGPGGSVDIVAMDEGSEASSCSSALASKPSPEGAS 420
              ||||||||||||||||||||||||||||
Db        361 SEHAQDTYLVLDKWLLPRNPPSEDLPG------------------------------- 387

Qy        421 AASFEYTILDPSSQLLRPWTLCPELPPTPHLKYLYLVVSDSGISTDYSSGDSQGAQGGL 480
              ||||||||||||||
Db        388 ------------PWALCPELPPTPHLKYLYLVVSDSGISTDYSSGDSQGAQGGL 430

Qy        481 SDGPYSNPYENSLIPAAEPLPPSYVACS 508
              |||||||||  ||||||||||||||
Db        431 SDGPYSSPYENSPIPAAEPLPPSYVACS 458
```

```
RESULT 12
ADM93420
ID   ADM93420 standard; protein; 458 AA.
XX
AC   ADM93420;
XX
DT   01-JUL-2004  (first entry)
XX
DE   Human NOVX polypeptide #26.
XX
KW   Human; NOVX; congenital heart defect; cardiomyopathy; atherosclerosis;
KW   hypertension; pulmonary stenosis; scleroderma; adenocarcinoma;
KW   haemophilia; graft-versus-host disease; cancer;
KW   neurodegenerative disorder; Alzheimer's disease; Parkinson's disease;
KW   multiple sclerosis; diabetes; obesity; bronchial asthma;
KW   acquired immunodeficiency syndrome; AIDS; Crohn's disease;
KW   infectious disease; anorexia; immune disorder.
XX
OS   Homo sapiens.
XX
PN   US2004067882-A1.
XX
PD   08-APR-2004.
XX
PF   05-NOV-2002; 2002US-00287971.
XX
PR   22-OCT-2001; 2001US-00035568.
PR   05-NOV-2001; 2001US-0338626P.
PR   06-NOV-2001; 2001US-0333072P.
PR   09-NOV-2001; 2001US-0345398P.
PR   09-NOV-2001; 2001US-0348283P.
PR   15-NOV-2001; 2001US-0335610P.
PR   21-NOV-2001; 2001US-0332152P.
PR   28-NOV-2001; 2001US-0333912P.
PR   29-NOV-2001; 2001US-00997425.
PR   29-NOV-2001; 2001US-0334100P.
PR   04-DEC-2001; 2001US-0365576P.
PR   05-FEB-2002; 2002US-0354807P.
PR   15-MAY-2002; 2002US-0380680P.
PR   16-MAY-2002; 2002US-0381043P.
PR   02-JUL-2002; 2002US-0393148P.
PR   02-JUL-2002; 2002US-0392262P.
PR   06-AUG-2002; 2002US-0401479P.
PR   06-AUG-2002; 2002US-0401626P.
PR   07-AUG-2002; 2002US-0401593P.
PR   07-AUG-2002; 2002US-0401695P.
PR   26-AUG-2002; 2002US-0406181P.
XX
PA   (ALSO/) ALSOBROOK J P.
PA   (ALVA/) ALVAREZ E.
PA   (ANDE/) ANDERSON D W.
PA   (BARO/) BARON M.
PA   (BOLD/) BOLDOG F L.
PA   (BURG/) BURGESS C E.
PA   (CASM/) CASMAN S J.
PA   (CHAP/) CHAPOVAL A.
PA   (DHAN/) DHANABAL M.
PA   (EDIN/) EDINGER S R.
PA   (EISE/) EISEN A.
PA   (ELLE/) ELLERMAN K.
PA   (ETTE/) ETTENBERG S.
PA   (GANG/) GANGOLLI E A.
PA   (GERL/) GERLACH V.
PA   (GORM/) GORMAN L.
PA   (GROS/) GROSSE W M.
PA   (GUOX/) GUO X.
PA   (HACK/) HACKETT C.
PA   (JIWW/) JI W.
PA   (KEKU/) KEKUDA R.
PA   (KHRA/) KHRAMTSOV N V.
PA   (LEPL/) LEPLEY D M.
PA   (LILL/) LI L.
PA   (MACD/) MACDOUGALL J R.
PA   (MALY/) MALYANKAR U M.
PA   (MAZU/) MAZUR A.
PA   (MCQU/) MCQUEENEY K.
PA   (MEZE/) MEZES P S.
PA   (MILL/) MILLER C E.
PA   (MILL/) MILLET I.
PA   (MISH/) MISHRA V.
PA   (PADI/) PADIGARU M.
PA   (PATT/) PATTURAJAN M.
PA   (PENA/) PENA C E A.
PA   (PEYM/) PEYMAN J A.
PA   (RAST/) RASTELLI L.
PA   (RIEG/) RIEGER D K.
PA   (ROTH/) ROTHENBERG M E.
PA   (SHEN/) SHENOY S G.
PA   (SHIM/) SHIMKETS R A.
PA   (SMIT/) SMITHSON G.
PA   (SPAD/) SPADERNA S K.
PA   (STAR/) STARLING G.
PA   (SPYT/) SPYTEK K A.
PA   (STON/) STONE D J.
PA   (TCHE/) TCHERNEV V T.
PA   (TWOM/) TWOMLOW N.
PA   (VERN/) VERNET C A M.
PA   (ZERH/) ZERHUSEN B D.
PA   (VOSS/) VOSS E Z.
PA   (ZHON/) ZHONG M.
XX
PI   Alsobrook JP, Alvarez E, Anderson DW, Baron M, Boldog FL;
PI   Burgess CE, Casman SJ, Chapoval A, Dhanabal M, Edinger SR, Eisen A;
PI   Ellerman K, Ettenberg S, Gangolli EA, Gerlach V, Gorman L;
PI   Grosse WM, Guo X, Hackett C, Ji W, Kekuda R, Khramtsov NV;
PI   Lepley DM, Li L, Macdougall JR, Malyankar UM, Mazur A, Mcqueeney K;
PI   Mezes PS, Miller CE, Millet I, Mishra V, Padigaru M, Patturajan M;
PI   Pena CEA, Peyman JA, Rastelli L, Rieger DK, Rothenberg ME;
PI   Shenoy SG, Shimkets RA, Smithson G, Spaderna SK, Starling G;
PI   Spytek KA, Stone DJ, Tchernev VT, Twomlow N, Vernet CAM;
PI   Zerhusen BD, Voss EZ, Zhong M;
XX
DR   WPI; 2004-355303/33.
DR   N-PSDB; ADM93419.
XX
PT   Novel isolated NOVX polypeptide useful treating or preventing disorders
PT   or syndromes such as Alzheimer's disease, Parkinson's disease, multiple
PT   sclerosis, diabetes, obesity, cancer, bronchial asthma, Crohn's disease.
XX
PS   Claim 2; SEQ ID NO 52; 330pp; English.
XX
CC   The invention relates to human NOVX polypeptides and the polynucleotides
```

CC encoding them. The NOVX polypeptides and polynucleotides are useful for
CC determining the presence of or predisposition to a disease associated
CC with altered levels of the sequences in a mammalian subject, and for
CC treating or preventing a pathology associated with NOVX. The
CC polypeptides, polynucleotides and antibodies that bind immunospecifically
CC to the polypeptides are useful for treating or preventing disorders or
CC syndromes such as congenital heart defects, cardiomyopathy,
CC atherosclerosis, hypertension, pulmonary stenosis, scleroderma,
CC adenocarcinoma, haemophilia, graft-versus-host disease, cancer,
CC neurodegenerative disorders, Alzheimer's disease, Parkinson's disease,
CC multiple sclerosis, diabetes, obesity, bronchial asthma, acquired
CC immunodeficiency syndrome (AIDS), Crohn's disease, infectious disease,
CC anorexia and immune disorders. This sequence represents a human NOVX
CC polypeptide of the invention. Note: The sequence data for this patent is
CC also available from USPTO at seqdata.uspto.gov/sequence.html.
XX
SQ Sequence 458 AA;

```
Query Match      88.7%;     Score 2407;  DB 8;   Length 458;
Best Local Similarity   89.2%;  Pred. No. 6.1e-191;
Matches   453;  Conservative   1;  Mismatches    4;  Indels    50;  Gaps    1;

Qy    1  MDHLGASLWPQVGSLCLLLAGAAWAPPNLPDPKFESKAALLAARGPEELLCFTERLEDL  60
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1  MDHLGASLWPQVGSLCLLLAGAAWAPPNLPDPKFESKAALLAARGPEELLCFTERLGDL  60

Qy   61  VCFWEEAASAGVGPGNYSFSYQLEDEPWKLCRLHQAPTARGAVRFWCSLPTADTSSFVPL 120
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61  VCFWEEAASAGVGPGNYSFSYQLEDEPWKLCRLHQAPTARGAVRFWCSLPTADTSSFVPL 120

Qy  121  ELRVTAASGAPRYHRVIHINEVVLLDAPVGLVARLADESGHVVLRWLPPPETPMTSHIRY 180
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  121  ELRVTAASGAPRYHRVIHINEVVLLDAPVGLVARLADESGHVVLRWLPPPETPMTSHIRY 180

Qy  181  EVDVSAGNGAGSVQRVEILEGRTECVLSNLRGRTRYTFAVRARMAEPSFGGFWSAWSEPV 240
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  181  EVDVSAGNGAGSVQRVEILEGRTECVLSNLRGRTRYTFAVRTRMAEPSFGGFWSAWSEPV 240

Qy  241  SLLTPSDLDPLILTLSLILVVILVLLTVLALLSHRRALKQKIWPGIPSPSEFEGLFTTH 300
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  241  SLLTPSDLDPLILTLSLILVVILVLLTVLALLSHRRALKQKIWPGIPSPSEFEGLFTTH 300

Qy  301  KGNFQLWLYQNDGCLWWSPCTPFTEDPPASLEVLSERCWGTMQAVEPGTDDEGPLLEPVG 360
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  301  KGNFQLWLYQNDGCLWWSPCTPFTEDPPASLEVLSERCWGTMQAVEPGTDDEGPLLEPVG 360

Qy  361  SEHAQDTYIVLDKWLLPRNPPSEDLPGPGGSVDIVAMDEGSEASCSSALASKPSPEGAS 420
         ||||||||||||||||||||||||||
Db  361  SEHAQDTYIVLDKWLLPRNPPSEDLPG-------------------------------- 387

Qy  421  AASFEYTILDPSSQLLRPWTLCPELPPTPPHLKLYLVVSDSGISTDYSSGDSQGAQGGL 480
               ||  ||||||||||||||||||||||||||||||||||||
Db  388  -------PWALCPELPPTPPHLKLYLVVSDSGISTDYSSGDSQGAQGGL 430

Qy  481  SDGPYSNPYENSLLIPAAEPLPPSVVACS 508
         |||||:||:|||:|||||||||||||||
Db  431  SDGPYSSPYENSPIPAAEPLPPSVVACS 458
```

RESULT 13
ADE28671
ID ADE28671 standard; protein; 458 AA.
XX
AC ADE28671;
XX
DT 29-JAN-2004 (first entry)
XX
DE Human NOV15a protein - SEQ ID 48.
XX
KW NOVX; antidiabetic; anorectic; cardiant; hypotensive;
KW antiarteriosclerotic; virucide; antibacterial; fungicide; protozoacide;
KW nootropic; neuroprotective; antiparkinsonian; anticonvulsant;
KW osteopathic; antiarthritic; antiinflammatory; dermatological;
KW antiasthmatic; antilipaemic; metabolic; diabetes; obesity; infectious;
KW anorexia; cancer; cardiovascular; hypertension; atherosclerosis;
KW neurodegenerative; Alzheimer's disease; Parkinson's; epilepsy; immune;
KW osteoarthritis; haemopoietic; inflammatory skin; asthma; dyslipidaemia;
KW neurogenesis; cell differentiation; proliferation; haemopoiesis;
KW wound healing; angiogenesis; gene therapy; chromosome mapping;
KW tissue typing; human; NOV.
XX
OS Homo sapiens.
XX
PN WO200304030-A2.
XX
PD 15-MAY-2003.
XX
PF 05-NOV-2002; 2002WO-US035536.
XX
PR 05-NOV-2001; 2001US-0336626P.
PR 05-DEC-2001; 2001US-0336600P.
PR 07-DEC-2001; 2001US-0338285P.
PR 12-DEC-2001; 2001US-0341346P.
PR 17-DEC-2001; 2001US-0341477P.
PR 17-DEC-2001; 2001US-0341540P.
PR 20-DEC-2001; 2001US-0342592P.
PR 27-DEC-2001; 2001US-0344297P.
PR 31-DEC-2001; 2001US-0344903P.
PR 17-APR-2002; 2002US-0373288P.
PR 15-MAY-2002; 2002US-0380981P.
PR 17-MAY-2002; 2002US-0381495P.
PR 28-MAY-2002; 2002US-0383534P.
PR 28-MAY-2002; 2002US-0383744P.
PR 29-MAY-2002; 2002US-0383829P.
PR 29-MAY-2002; 2002US-0384024P.
PR 07-AUG-2002; 2002US-0401788P.
PR 26-AUG-2002; 2002US-0406353P.
PR 31-OCT-2002; 2002US-00287971.
XX
PA (CURA-) CURAGEN CORP.
XX
PI Alsobrook JP, Alvarez E, Anderson DW, Baron M, Boldog FL;
PI Burgess CE, Casman SJ, Chapoval A, Dhanabal M, Edinger SR, Eisen A;
PI Ellerman K, Ettenberg S, Gangolli EA, Gerlach VL, Gorman L;
PI Grosse WM, Guo X, Hackett C, Ji H, Kekuda R, Khramtsov NV;
PI Lepley DM, Li L, Macdougall JR, Millet I, Malyankar UM, Mazur A, Mcqueeney K;
PI Mezes PS, Miller CE, Millet I, Mishra VS, Padigaru M, Patturajan M;
PI Pena CEA, Peyman JA, Rastelli L, Rieger DK, Shenoy SG, Shimkets RA;
PI Smithson G, Starling G, Spytek KA, Stone DJ, Tcherneva VT, Twomlow N;
PI Vernet CAM, Zerhusen BD, Zhong M;
XX
DR WPI; 2003-441555/41.
DR N-PSDB; ADE28670.
XX
PT New isolated NOVX polypeptides and polynucleotides, useful for
PT preventing, diagnosing or treating NOVX-associated disorders, e.g.
PT osteoarthritis, obesity, atherosclerosis, cancer, Parkinson's disease,
PT asthma, or infections.
XX
PS Claim 1; SEQ ID NO 48; 447pp; English.
XX
CC The invention relates to a novel isolated NOVX polypeptide. The
CC polypeptide of the invention demonstrates, antidiabetic, anorectic,
CC cardiant, hypotensive, antiarteriosclerotic, virucide, antibacterial,
CC fungicide, protozoacide, nootropic, neuroprotective, antiparkinsonian,
CC anticonvulsant, osteopathic, antiarthritic, antiinflammatory,
CC dermatological, antiasthmatic and antilipaemic activities. The
CC polypeptides, nucleic acid molecules and antibodies may be useful for
CC treating or diagnosing diseases including metabolic disorders such as
CC diabetes and obesity, infectious diseases, anorexia, cancer,
CC cardiovascular diseases including hypertension and atherosclerosis,
CC neurodegenerative disorders such as Alzheimer's disease, Parkinson's
CC disease and epilepsy, immune disorders e.g. osteoarthritis, haemopoietic
CC disorders, inflammatory skin disorders, asthma and dyslipidaemia.
CC Furthermore, the nucleic acids and polypeptides may also be used to
CC identify molecules that modulate or inhibit neurogenesis, cell

```
CC    differentiation and proliferation, haemopoiesis, wound healing and
CC    angiogenesis, as well as in gene therapy. Finally, the nucleic acids may
CC    be used as hybridisation probes, in chromosome mapping, tissue typing,
CC    preventive medicine and pharmacogenomics. The current sequence is that of
CC    the human NOV protein of the invention.
XX
SQ    Sequence 458 AA;

Query Match      88.5%;   Score 2404;  DB 7;  Length 458;
Best Local Similarity  89.2%;   Pred. No. 1.1e-190;
Matches  453;  Conservative  .1; Mismatches  4; Indels  .1; Gaps  50;  Gaps  1;

Qy       1  MDHLGASLWPQVGSLCLLLAGAAWAPPPNLPDPKFESKAALLAARGPEELLCFTERLEDL   60
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1  MDHLGASLWPQVGSLCLLLAGAAWAPPPNLPDPKFESKAALLAARGPEELLCFTERLEDL   60

Qy      61  VCFWEEAASAGVGPGNYSFSYQLEDEPWKLCRLHQAPTARGAVRFWCSLPTADTSSFVPL  120
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      61  VCFWEEAASAGVGPGNYSFSYQLEDEPWKLCRLHQAPTARGAVRFWCSLPTADTSSFVPL  120

Qy     121  ELRVTAASGAPRYHRVIHINEVVLLDAPVGLVARLADESGHVVLRWLPPETPMTSHIRY  180
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     121  ELRVTAASGAPRYHRVIHINEVVLLDAPVGLVARLADESGHVVLRWLPPETPMTSHIRY  180

Qy     181  EVDVSAGNGAGSVQRVEILEGRTECVLSNLRGRTRYTFAVRARMAEPSFGGFWSAWSEPV  240
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     181  EVDVSAGNGAGSVQRVEILEGRTECVLSNLRGRTRYTFAVRARMAEPSFGGFWSAWSEPV  240

Qy     241  SLLTPSDLDPLIITLSLILVVILVVLLTVLALLSHRRALKQKIWPGIPSPESEFEGLFTTH  300
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     241  SLLTPSDLDPLIITLSLILVVILVVLLTVLALLSHRRALKQKIWPGIPSPESEFEGLFTTH  300

Qy     301  KGNFQLWLYQNDGCLWWSPCTPFTEDDPPASLEVLSERCWGTMQAVEPGTDDBGPLLEPVG  360
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     301  KGNFQLWLYQNDGCLWWSACTPFTEDDPAFLEVLSERCWGTMQAVEPGTDDEGPLLEPVG  360

Qy     361  SEHAQDTYLVLDKWLLPRNPPSEDLPGPGGSVDIVAMDEGSEASSCSSALASKPSPEGAS  420
            |||||||||||||||||||||||||||
Db     361  SEHAQDTYLVLDKWLLPRNPPSEDLPG------------------------------  387

Qy     421  AASFEYTILDPSSQLLRPWTLCPELPTPPHLKYLYLVVSDSGISTDYSSGDSQGAQGGL  480
                                      ||||||||||||||||||||||||||||||||
Db     388  ------------------------PWALCPELPTPPHLKYLYLVVSDSGISTDYSSGDSQGAQGGL  430

Qy     481  SDGPYSNPYENSLIPAAEPLPPSYVACS  508
            ||||||:|||||||||||||||
Db     431  SDGPYSSPYENSPIPAAEPLPPSYVACS  458


RESULT 14
ADE28673
ID    ADE28673 standard; protein; 458 AA.
XX
AC    ADE28673;
XX
DT    29-JAN-2004 (first entry)
XX
DE    Human NOV15b protein - SEQ ID 50.
XX
KW    NOVX; antidiabetic; anorectic; cardiant; hypotensive;
KW    antiarteriosclerotic; virucide; antibacterial; fungicide; protozoacide;
KW    nootropic; neuroprotective; antiparkinsonian; anticonvulsant;
KW    osteopathic; antiarthritic; antiinflammatory; dermatological;
KW    antiasthmatic; antilipaemic; metabolic; diabetes; obesity; infectious;
KW    anorexia; cancer; cardiovascular; hypertension; atherosclerosis;
KW    neurodegenerative; Alzheimer's disease; Parkinson's; epilepsy; immune;
KW    osteoarthritis; haemopoietic; inflammatory skin; asthma; dyslipidaemia;
KW    neurogenesis; cell differentiation; proliferation; haemopoiesis;
KW    wound healing; angiogenesis; gene therapy; chromosome mapping;
KW    tissue typing; human; NOV.
XX
OS    Homo sapiens.
XX
```

```
PN    WO200304030-A2.
XX
PD    15-MAY-2003.
XX
PF    05-NOV-2002; 2002WO-US035536.
XX
PR    05-NOV-2001; 2001US-0338626P.
PR    05-DEC-2001; 2001US-0336600P.
PR    07-DEC-2001; 2001US-0338285P.
PR    12-DEC-2001; 2001US-0341346P.
PR    17-DEC-2001; 2001US-0341477P.
PR    17-DEC-2001; 2001US-0341540P.
PR    20-DEC-2001; 2001US-0342592P.
PR    27-DEC-2001; 2001US-0344297P.
PR    31-DEC-2001; 2001US-0344903P.
PR    17-APR-2002; 2002US-0373288P.
PR    15-MAY-2002; 2002US-0380981P.
PR    17-MAY-2002; 2002US-0381495P.
PR    28-MAY-2002; 2002US-0383534P.
PR    28-MAY-2002; 2002US-0383744P.
PR    29-MAY-2002; 2002US-0383829P.
PR    29-MAY-2002; 2002US-0384024P.
PR    07-AUG-2002; 2002US-0401788P.
PR    26-AUG-2002; 2002US-0406353P.
PR    31-OCT-2002; 2002US-00287971.
XX
PA    (CURA-) CURAGEN CORP.
XX
PI    Alsobrook JP,  Alvarez E,  Anderson DW,  Baron M,  Boldog FL;
PI    Burgess CE,  Casman SJ,  Chapoval A,  Dhanabal M,  Edinger SR,  Eisen A;
PI    Ellerman K,  Ettenberg S,  Gangolli EA,  Gerlach VL,  Gorman L;
PI    Grosse WM,  Guo X,  Hackett C,  Ji W,  Kekuda R,  Khramtsov NV;
PI    Lepley DM,  Li L,  Macdougall JR,  Malyankar UM,  Mazur A,  Mcqueeney K;
PI    Mezes PS,  Miller CE,  Millet I,  Mishra VS,  Padigaru M,  Patturajan M;
PI    Pena CEA,  Peyman JA,  Rastelli L,  Rieger DK,  Shenoy SG,  Shimkets RA;
PI    Smithson G,  Starling G,  Spytek KA,  Stone DJ,  Tchernev VT,  Twomlow N;
PI    Vernet CAM,  Zerhusen BD,  Zhong M;
XX
DR    WPI; 2003-441555/41.
DR    N-PSDB; ADE28672.
XX
PT    New isolated NOVX polypeptides and polynucleotides, useful for
PT    preventing, diagnosing or treating NOVX-associated disorders, e.g.
PT    osteoarthritis, obesity, atherosclerosis, cancer, Parkinson's disease,
PT    asthma, or infections.
XX
PS    Claim 1; SEQ ID NO 50; 447pp; English.
XX
CC    The invention relates to a novel isolated NOVX polypeptide. The
CC    polypeptide of the invention demonstrates, antidiabetic, anorectic,
CC    cardiant, hypotensive, antiarteriosclerotic, virucide, antibacterial,
CC    fungicide, protozoacide, nootropic, neuroprotective, antiparkinsonian,
CC    anticonvulsant, osteopathic, antiarthritic, antiinflammatory,
CC    dermatological, antiasthmatic and antilipaemic activities. The
CC    polypeptides, nucleic acid molecules and antibodies may be useful for
CC    treating or diagnosing diseases including metabolic disorders such as
CC    diabetes and obesity, infectious diseases, anorexia, cancer,
CC    cardiovascular diseases including hypertension and atherosclerosis,
CC    neurodegenerative disorders such as Alzheimer's disease, Parkinson's
CC    disease and epilepsy, immune disorders e.g. osteoarthritis, haemopoietic
CC    disorders, inflammatory skin disorders, asthma and dyslipidaemia.
CC    Furthermore, the nucleic acids and polypeptides may also be used to
CC    identify molecules that modulate or inhibit neurogenesis, cell
CC    differentiation and proliferation, haemopoiesis, wound healing and
CC    angiogenesis, as well as in gene therapy. Finally, the nucleic acids may
CC    be used as hybridisation probes, in chromosome mapping, tissue typing,
CC    preventive medicine and pharmacogenomics. The current sequence is that of
CC    the human NOV protein of the invention.
XX
SQ    Sequence 458 AA;

Query Match      88.5%;       Score 2404;  DB 7;  Length 458;
Best Local Similarity  89.2%;  Pred. No. 1.1e-190;
```

```
Matches  453;  Conservative  1;  Mismatches  4;  Indels  50;  Gaps  1;

Qy    1 MDHLGASLWPQVGSLCLLLAGAAWAPPPNLPDPKFESKAALLAARGPEELLCFTERLEDL 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1 MDHLGASLWPQVGSLCLLLAGAAWAPPPNLPDPKFESKAALLAARGPEELLCFTERLEDL 60

Qy   61 VCFWEEAASAGVGPGNYSFSYQLEDEPWKLCRLHQAPTARGAVRFWCSLPTADTSSFVPL 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61 VCFWEEAASAGVGPGNYSFSYQLEDEPWKLCRLHQAPTARGAVRFWCSLPTADTSSFVPL 120

Qy  121 ELRVTAASGAPRYHRVIHINEVVLLDAPVGLVARLADESGHVVLRWLPPETPMTSHIRY 180
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  121 ELRVTAASGAPRYHRVIHINEVVLLDAPVGLVARLADESGHVVLRWLPPETPMTSHIRY 180

Qy  181 EVDVSAGNGAGSVQRVEILEGRTECVLSNLRGRTRYTFAVRARMAEPSFGGFWSAWSEPV 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  181 EVDVSAGNGAGSVQRVEILEGRTECVLSNLRGRTRYTFAVRARMAEPSFGGFWSAWSEPV 240

Qy  241 SLLTPSDLDPLIITLSLILVVILVLLTVLALLSHRRALKQKIWPGIPSPESEFEGLFTTH 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  241 SLLTPSDLDPLIITLSLILVVILVLLTVLALLSHRRALKQKIWPGIPSPESEFEGLFTTH 300

Qy  301 KGNFQLWLYQNDGCLWWSPCTPFTEDDPASLEVLSERCWGTMQAVEPGTDDEGPLLEPVG 360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  301 KGNFQLWLYQNDGCLWWSACTPFTEDDPAFLEVLSERCWGTMQAVEPGTDDEGPLLEPVG 360

Qy  361 SEHAQDTYLVLDKWLLPRNPPSEDLPGPGGSVDIVAMDEGSEASSCSSALASKPSPEGAS 420
          ||||||||||||||||||||||||||
Db  361 SEHAQDTYLVLDKWLLPRNPPSEDLPG------------------------------- 387

Qy  421 AASFEYTILDPSSQLLRPWTLCPELPPTPPHLKYLYLVVSDSGISTDYSSGDSQGAQGGL 480
                        |||||||||||||||||||||||||||||||
Db  388 --------------PWALCPELPPTPPHLKYLYLVVSDSGISTDYSSGDSQGAQGGL 430

Qy  481 SDGPYSNPYENSLIPAAEPLPPSYVACS 508
          |||||||:||||||||||||||||||||
Db  431 SDGPYSSPYENSPIPAAEPLPPSYVACS 458
```

RESULT 15
ADM93416
ID   ADM93416 standard; protein; 458 AA.
XX
AC   ADM93416;
XX
DT   01-JUL-2004 (first entry)
XX
DE   Human NOVX polypeptide #24.
XX
KW   Human; NOVX; congenital heart defect; cardiomyopathy; atherosclerosis;
KW   hypertension; pulmonary stenosis; scleroderma; adenocarcinoma;
KW   haemophilia; graft-versus-host disease; cancer;
KW   neurodegenerative disorder; Alzheimer's disease; Parkinson's disease;
KW   multiple sclerosis; diabetes; obesity; bronchial asthma;
KW   acquired immunodeficiency syndrome; AIDS; Crohn's disease;
KW   infectious disease; anorexia; immune disorder.
XX
OS   Homo sapiens.
XX
PN   US2004067882-A1.
XX
PD   08-APR-2004.
XX
PF   05-NOV-2002; 2002US-00287971.
XX
PR   22-OCT-2001; 2001US-00035568.
PR   05-NOV-2001; 2001US-0338626P.
PR   06-NOV-2001; 2001US-0333072P.
PR   09-NOV-2001; 2001US-0345398P.
PR   09-NOV-2001; 2001US-0348283P.
PR   15-NOV-2001; 2001US-0335610P.
PR   21-NOV-2001; 2001US-0332152P.
PR   28-NOV-2001; 2001US-0333912P.
PR   29-NOV-2001; 2001US-00997425.
PR   29-NOV-2001; 2001US-0334300P.
PR   04-DEC-2001; 2001US-0336576P.
PR   05-FEB-2002; 2002US-0354807P.
PR   15-MAY-2002; 2002US-0380968P.
PR   16-MAY-2002; 2002US-0381043P.
PR   02-JUL-2002; 2002US-0393148P.
PR   02-JUL-2002; 2002US-0393262P.
PR   06-AUG-2002; 2002US-0401479P.
PR   06-AUG-2002; 2002US-0401626P.
PR   07-AUG-2002; 2002US-0401593P.
PR   07-AUG-2002; 2002US-0401695P.
PR   26-AUG-2002; 2002US-0406181P.
XX
PA   (ALSO/) ALSOBROOK J P.
PA   (ALVA/) ALVAREZ E.
PA   (ANDE/) ANDERSON D W.
PA   (BARO/) BARON M.
PA   (BOLD/) BOLDOG F L.
PA   (BURG/) BURGESS C E.
PA   (CASM/) CASMAN S J.
PA   (CHAP/) CHAPOVAL A.
PA   (DHAN/) DHANABAL M.
PA   (EDIN/) EDINGER S R.
PA   (EISE/) EISEN A.
PA   (ELLE/) ELLERMAN K.
PA   (ETTE/) ETTENBERG S.
PA   (GANG/) GANGOLLI E A.
PA   (GERL/) GERLACH V.
PA   (GORM/) GORMAN L.
PA   (GROS/) GROSSE W M.
PA   (GUOX/) GUO X.
PA   (HACK/) HACKETT C.
PA   (JIWW/) JI W.
PA   (KEKU/) KEKUDA R.
PA   (KHRA/) KHRAMTSOV N V.
PA   (LEPL/) LEPLEY D M.
PA   (LILL/) LI L.
PA   (MACD/) MACDOUGALL J R.
PA   (MALY/) MALYANKAR U M.
PA   (MAZU/) MAZUR A.
PA   (MCQU/) MCQUEENEY K.
PA   (MEZE/) MEZES P S.
PA   (MILL/) MILLER C E.
PA   (MILL/) MILLET I.
PA   (MISH/) MISHRA V.
PA   (PADI/) PADIGARU M.
PA   (PATT/) PATTURAJAN M.
PA   (PENA/) PENA C E A.
PA   (PEYM/) PEYMAN J A.
PA   (RAST/) RASTELLI L.
PA   (RIEG/) RIEGER D K.
PA   (ROTH/) ROTHENBERG M E.
PA   (SHEN/) SHENOY S G.
PA   (SHIM/) SHIMKETS R A.
PA   (SMIT/) SMITHSON G.
PA   (SPAD/) SPADERNA S K.
PA   (STAR/) STARLING G.
PA   (SPYT/) SPYTEK K A.
PA   (STON/) STONE D J.
PA   (TCHE/) TCHERNEV V T.
PA   (TWOM/) TWOMLOW N.
PA   (VERN/) VERNET C A M.
PA   (ZERH/) ZERHUSEN B D.
PA   (VOSS/) VOSS E Z.
PA   (ZHON/) ZHONG M.
XX
PI   Alsobrook JP, Alvarez E, Anderson DW, Baron M, Boldog FL;
PI   Burgess CE, Casman SJ, Chapoval A, Dhanabal M, Edinger SR, Eisen A;
PI   Ellerman K, Ettenberg S, Gangolli EA, Gerlach V, Gorman L;
PI   Grosse WM, Guo X, Hackett C, Ji W, Kekuda R, Khramtsov NV;
PI   Lepley DM, Li L, Macdougall JR, Malyankar UM, Mazur A, Mcqueeney K;
PI   Mezes PS, Miller CE, Millet I, Mishra V, Padigaru M, Patturajan M;

Search completed: April 21, 2005, 08:51:50
Job time : 191 secs

PI   Pena CEA,   Peyman JA,   Rastelli L,   Rieger DK,   Rothenberg ME;
PI   Shenoy SG,   Shimkets RA,   Smithson G,   Spaderna SK,   Starling G;
PI   Spytek KA,   Stone DJ,   Tchernev VT,   Twomlow N,   Vernet CAM;
PI   Zerhusen BD,   Voss EZ,   Zhong M;
XX
DR   WPI; 2004-355303/33.
DR   N-PSDB; ADM93415.
XX
PT   Novel isolated NOVX polypeptide useful treating or preventing disorders
PT   or syndromes such as Alzheimer's disease, Parkinson's disease, multiple
PT   sclerosis, diabetes, obesity, cancer, bronchial asthma, Crohn's disease.
XX
PS   Claim 2; SEQ ID NO 48; 330pp; English.
XX
CC   The invention relates to human NOVX polypeptides and the polynucleotides
CC   encoding them. The NOVX polypeptides and polynucleotides are useful for
CC   determining the presence of or predisposition to a disease associated
CC   with altered levels of the sequences in a mammalian subject, and for
CC   treating or preventing a pathology associated with NOVX. The
CC   polypeptides, polynucleotides and antibodies that bind immunospecifically
CC   to the polypeptides are useful for treating or preventing disorders or
CC   syndromes such as congenital heart defects, cardiomyopathy,
CC   atherosclerosis, hypertension, pulmonary stenosis, scleroderma,
CC   adenocarcinoma, haemophilia, graft-versus-host disease, cancer,
CC   neurodegenerative disorders, Alzheimer's disease, Parkinson's disease,
CC   multiple sclerosis, diabetes, obesity, bronchial asthma, acquired
CC   immunodeficiency syndrome (AIDS), Crohn's disease, infectious disease,
CC   anorexia and immune disorders. This sequence represents a human NOVX
CC   polypeptide of the invention. Note: The sequence data for this patent is
CC   also available from USPTO at seqdata.uspto.gov/sequence.html.
XX
SQ   Sequence 458 AA;

Query Match   88.5%;   Score 2404;   DB 8;   Length 458;
Best Local Similarity   89.2%;   Pred. No. 1.1e-190;
Matches 453;   Conservative   1;   Mismatches   4;   Indels 50;   Gaps   1;

Qy   1 MDHLGASLWPQVGSLCLLLAGAAWAPPPNLPDPKFESKAALLAARGPEELLCFTERLEDL   60
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1 MDHLGASLWPQVGSLCLLLAGAAWAPPPNLPDPKFESKAALLAARGPEELLCFTERLEDL   60

Qy   61 VCFWEEAASAGVGPGNYSFSYQLEDEPWKLCRLHQAPTARGAVRFWCSLPTADTSSFVPL 120
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61 VCFWEEAASAGVGPGNYSFSYQLEDEPWKLCRLHQAPTARGAVRFWCSLPTADTSSFVPL 120

Qy   121 ELRVTAASGAPRYHRVIHINEVVLLDAPVGLVARLADESGHVVLRWLPPPETPMTSHIRY 180
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   121 ELRVTAASGAPRYHRVIHINEVVLLDAPVGLVARLADESGHVVLRWLPPPETPMTSHIRY 180

Qy   181 EVDVSAGNGAGSVQRVEILEGRTECVLSNLRGRTRYTFAVRARMAEPSFGGFWSAWSEPV 240
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   181 EVDVSAGNGAGSVQRVEILEGRTECVLSNLRGRTRYTFAVRARMAEPSFGGFWSAWSEPV 240

Qy   241 SLLTPSDLDPLILTLSLILVVILVVILVLLTVLALLSHRRALKQKIWPGIPSPSEFEGLFTTH 300
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   241 SLLTPSDLDPLILTLSLILVVILVVILVLLTVLALLSHRRALKQKIWPGIPSPSEFEGLFTTH 300

Qy   301 KGNFQLWLYQNDGCLWWSPCTPFTEDPPASLEVLSERCWGTMQAVEPGTDDEGPLLEPVG 360
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   301 KGNFQLWLYQNDGCLWWSACTPFTEDPPAFLEVLSERCWGTMQAVEPGTDDEGPLLEPVG 360

Qy   361 SEHAQDTYLVLDKWLLPRNPPSEDLPGPGGSVDIVAMDEGSEASSCSSALASKPSPEGAS 420
            ||||||||||||||||||||||||||||||||----------------------------
Db   361 SEHAQDTYLVLDKWLLPRNPPSEDLPG--------------------------------- 387

Qy   421 AASFEYTILDPSSQLLRPWTLCPELPPTPHLKYLYLVVSDSGISTDYSSGDSQGAQGGL   480
            ||||||||||||||||||||||||||||||||||||||||||||
Db   388 ------------PWALCPELPPTPHLKYLYLVVSDSGISTDYSSGDSQGAQGGL   430

Qy   481 SDGPYSNPYENSLIPAAEPLPPSYVACS 508
            ||||:::|||||:||:||||||||||||
Db   431 SDGPYSSPYENSPIPAAEPLPPSYVACS 458

```
                    GenCore version 5.1.6
         Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on:      April 20, 2005, 21:00:26 ; Search time 5262 Seconds
                                        (without alignments)
                                        11046.017 Million cell updates/sec

Title:       US-09-016-159D-4
Perfect score: 1527
Sequence:    1 atggaccacctcggggcgtc.........gctatgtggcttgctcttag 1527

Scoring table: IDENTITY_NUC
             Gapop 10.0 , Gapext 1.0

Searched:    34239544 seqs, 19032134700 residues

Total number of hits satisfying chosen parameters:    68479088

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :   EST:*
             1: gb_est1:*
             2: gb_est2:*
             3: gb_htc:*
             4: gb_est3:*
             5: gb_est4:*
             6: gb_est5:*
             7: gb_est6:*
             8: gb_gss1:*
             9: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

                          SUMMARIES
```

| | | | % | | | | |
Result	Score	Query Match	Length	DB	ID		Description
No.							
1	1527	100.0	1759	3	CR592865	CR592865 full-leng
2	1527	100.0	1797	3	CR620075	CR620075 full-leng
3	1527	100.0	1811	9	CR618473	CR618473 full-leng
4	1455	95.3	1527	9	AY414846	AY414846 Homo sapi
5	1413	92.5	1870	3	CR610088	CR610088 full-leng
6	1271.6	83.3	1681	9	CR613702	CR613702 full-leng
7	1005.8	65.9	1452	9	AY414848	AY414848 Mus muscu
c 8	977	64.0	1082	5	BX382170	BX382170 Mus muscu
9	949.6	62.2	1524	3	AK010968	AK010968 Mus muscu
10	943	61.8	1320	2	CR597800	CR597800 full-leng
11	863.8	56.6	984	5	BX448344	BX448344 BX448344
12	857.4	56.1	1079	1	AL553257	AL553257 AL553257
13	812.2	53.2	1005	5	BX335579	BX335579 BX335579
c 14	792.2	51.9	1051	5	BX448343	BX448343 BX448343
15	789	51.7	1115	9	AY414847	AY414847 Pan trogl
16	773.4	50.6	775	5	BX096321	BX096321 BX096321
c 17	762.4	49.9	968	5	BX406157	BX406157 BX406157
18	749.8	49.1	997	3	AL523468	AL523468 AL523468
c 19	747.4	48.9	1075	1	AL576311	AL576311 AL576311
c 20	745.6	48.8	1127	1	AL523894	AL523894 AL523894
21	722.4	47.3	918	5	BX378570	BX378570 BX378570
c 22	693.8	45.4	881	1	AL521907	AL521907 AL521907
c 23	687	45.0	769	4	BM020450	BM020450 603649132
c 24	678	44.4	972	1	AL523467	AL523467 AL523467
25	672	44.0		3	1231	CR592320	CR592320 full-leng AGENCOURT
26	669.4	43.8		5	994	BQ894763	BQ894763 AGENCOURT
27	653.4	42.8		5	1051	AL523895	AL523895 AL523895
28	640	41.9		5	1030	BX340440	BX340440 BX340440
29	634.2	41.5		4	640	BI760304	BI760304 603045374
30	630	41.3		1	902	AL523633	AL523633 AL523633
31	610.4	40.0		5	816	BX406158	BX406158 BX406158
c 32	604.4	39.6		5	973	BX335578	BX335578 BX335578
c 33	561.8	36.8		1	945	AL523632	AL523632 AL523632
c 34	559.4	36.6		1	944	AL521906	AL521906 AL521906
35	558.4	36.6		6	1009	BY710355	BY710355 BY710355
c 36	544	35.6		6	1098	BX340439	BX340439 BX340439
37	541.6	35.5		5	1054	BX382171	BX382171 BX382171
38	509.6	33.4		5	890	BG105276	BG105276 602312859
39	501.6	32.8		6	951	CD518752	CD518752 AGENCOURT
40	487.8	31.9		5	853	BQ919762	BQ919762 AGENCOURT
c 41	478.4	31.3		5	724	BQ773300	BQ773300 UI-H-FE0-
42	470	30.8		6	488	CB269218	CB269218 1008125 H
43	449.2	29.4		6	1006	BI838739	BI838739 603082183
44	440	28.8		4	984	BG163892	BG163892 60343252
45	419	27.4		2	534	BF078672	BF078672 229067 MA

```
                          ALIGNMENTS

RESULT 1
CR592865
LOCUS       CR592865             1759 bp   mRNA    linear   HTC 21-JUL-2004
DEFINITION  full-length cDNA clone CSODI073YP22 of Placenta Cot 25-normalized
            of Homo sapiens (human).
ACCESSION   CR592865
VERSION     CR592865.1  GI:50473672
KEYWORDS    HTC; CNSLT cDNA.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 1759)
  AUTHORS   Li,W.B., Gruber,C., Jessee,J. and Polayes,D.
  TITLE     Full-length cDNA libraries and normalization
  JOURNAL   Unpublished
  REMARK    Contact : Feng Liang Email : fliang@lifetech.com URL :
            http://fulllength.invitrogen.com/ InvitroGen Corporation 1600
            Faraday Avenue
REFERENCE   2  (bases 1 to 1759)
  AUTHORS   Genoscope.
  TITLE     Direct Submission
  JOURNAL   Submitted (20-JUL-2004) Genoscope - Centre National de Sequencage ;
            BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr)
            - Web : www.genoscope.cns.fr)
COMMENT     1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
            end enriched, double-strand cDNA was digested with Not I and cloned
            into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library
            was normalized. Library was constructed by Life Technologies, a
            division of Invitrogen.
FEATURES             Location/Qualifiers
     source          1..1759
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="CSODI073YP22"
                     /tissue_type="Placenta Cot 25-normalized"
                     /plasmid="pCMVSPORT_6"
ORIGIN
Query Match          100.0%;  Score 1527;  DB 3;  Length 1759;
Best Local Similarity 100.0%;  Pred. No. 0;
Matches 1527;  Conservative  0;  Mismatches  0;  Indels  0;  Gaps  0;

Qy    1 ATGGACCACCTCGGGGCGTCCCCTCTGGCCCCAGGTCGGCTCCCTTGTCTCCTGCTCGCT 60
Db   44 ATGGACCACCTCGGGGCGTCCCCTCTGGCCCCAGGTCGGCTCCCTTTGTCTCCTGCTCGCT 103
```

```
Qy    61 GGGGCCGCCTGGGGCGCCCCCGCTAACCTCCCGGACCCCAAGTTGAGAGCAAAGCGGCC  120
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   104 GGGGCCGCCTGGGGCGCCCCCGCTAACCTCCCGGACCCCAAGTTGAGAGCAAAGCGGCC  163

Qy   121 TTGCTGGCGGCCCGGGGCCCGAGAGAGTTCTGTGCTTCACCGAGCGCGGTTGAGGGACTTG  180
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   164 TTGCTGGCGGCCCGGGGCCCGAGAGAGTTCTGTGCTTCACCGAGCGCGGTTGAGGGACTTG  223

Qy   181 GTGTGTTTCGGGAGGAAGCGGCGAGGGCTGGGGTGGCCCGGGCAACTACAGCTTCTCC  240
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   224 GTGTGTTTCGGGAGGAAGCGGCGAGGGCTGGGGTGGCCCGGGCAACTACAGCTTCTCC  283

Qy   241 TACCAGTCTGAGGATGAGCCATGGAAGCTGTGTCGCCTGCACCAGGCTCCCACGGCTCGT  300
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   284 TACCAGTCTGAGGATGAGCCATGGAAGCTGTGTCGCCTGCACCAGGCTCCCACGGCTCGT  343

Qy   301 GGTGCGGTGCGCTTCGTGTTCGCTCCTACAGCCGACACGTCGAGCTTCGTGCCCCTA  360
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   344 GGTGCGGTGCGCTTCGTGTTCGCTCCTACAGCCGACACGTCGAGCTTCGTGCCCCTA  403

Qy   361 GAGTTGCGCGTCACAGCAGCCTTCGGCGGCGTCCGGGCGCTCCGGGATATCACCGTGTCATCCACATCAAT  420
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   404 GAGTTGCGCGTCACAGCAGCCTTCGGCGGCGTCCGGGCGCTCCGGGATATCACCGTGTCATCCACATCAAT  463

Qy   421 GAAGTAGTGCTCCTAGACGCCCCCGTGGGGCTGGTGGCGGTTGGCTGACGAGAGCGGC  480
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   464 GAAGTAGTGCTCCTAGACGCCCCCGTGGGGCTGGTGGCGGTTGGCTGACGAGAGCGGC  523

Qy   481 CACGTAGTGTTGCGCTGGCTGCTCCCGCCCTGAGACACGTCTCACATCCGCTAC  540
         |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   524 CACGTAGTGTTGCGCTGGCTGCTCCCGCCCTGAGACACGTCTCACATCCGCTAC  583

Qy   541 GAGGTGACGTCTCGGCCCGGCAACGGCGCAGGGAGCCTACAGAGAGGTGGAGATCCTGGAG  600
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   584 GAGGTGACGTCTCGGCCCGGCAACGGCGCAGGGAGCCTACAGAGAGGTGGAGATCCTGGAG  643

Qy   601 GGCCGCACCGAGTGTGTGCTGAGCGACAACCTGCGCGGGCCCGGACGCGCGTACACCTTCGCCGTC  660
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   644 GGCCGCACCGAGTGTGTGCTGAGCGACAACCTGCGCGGGCCCGGACGCGCGTACACCTTCGCCGTC  703

Qy   661 CGGCGCGGTATGCGTGACCGCTTCGGCAGCTTCTGGAGCGCCTGTGCCGAGCCGTG  720
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   704 CGGCGCGGTATGCGTGACCGCTTCGGCAGCTTCTGGAGCGCCTGTGCCGAGCCGTG  763

Qy   721 TCGCTGCTGAGCGACCTGAGCCGACCCCCCATCGCTGACGCTCCTCCTCATCCTCGTG  780
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   764 TCGCTGCTGAGCGACCTGAGCCGACCCCCCATCGCTGACGCTCCTCCTCATCCTCGTG  823

Qy   781 GTCATCCTCGGTGTCGCTGACCGTGCGTCGCTCCGCGGGCTCTCGAAGCAG  840
         |||||||||||||||||||||||||||||||||||||||||||||||||||
Db   824 GTCATCCTCGGTGTCGCTGACCGTGCGTCGCTCCGCGGGCTCTCGAAGCAG  883

Qy   841 AAGATCTGGCCTGGCATCCCAGAGAGCGAGTTTGAAGGCCTTCACCACCAC  900
         ||||||||||||||||||||||||||||||||||||||||||||||||||
Db   884 AAGATCTGGCCTGGCATCCCAGAGAGCGAGTTTGAAGGCCTTCACCACCAC  943

Qy   901 AAGGGTAACTTCCAGCTGTGGCTGTACCAGAATGATGGCTGCTGCTTCCTGGTGGTGTGGAGCCCCTGC  960
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   944 AAGGGTAACTTCCAGCTGTGGCTGTACCAGAATGATGGCTGCTGCTTCCTGGTGGTGTGGAGCCCCTGC 1003

Qy   961 ACCCCCTTCACGGAGCGAGGACCCCACCTGCTTCCTCGGGGG 1020
         ||||||||||||||||||||||||||||||||||||||||||
Db  1004 ACCCCCTTCACGGAGCGAGGACCCCACCTGCTTCCTCGGGGG 1063

Qy  1021 ACGATGCAGGCAGTGAGCGCCCGGGACAGATGATGAGGGCCAGTGGGC 1080
         ||||||||||||||||||||||||||||||||||||||||||||||||
Db  1064 ACGATGCAGGCAGTGAGCGCCCGGGACAGATGATGAGGGCCAGTGGGC 1123

Qy  1081 AGTGAGCATGCCCAGATACCTATCTGGTGCTGGACAAATGTTGCTGCCCCGGAACCCG 1140
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1124 AGTGAGCATGCCCAGGATACCTATCTGTGCTGGACAAATGTTGCTGCTGCCCCGGAACCCG 1183

Qy  1141 CCCAGTGAGGACCTCCCAGGGCCTGGTGCGCAGTGTGGACATAGTGGCCATGATGAAGGC 1200
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1184 CCCAGTGAGGACCTCCCAGGGCCTGGTGCGCAGTGTGGACATAGTGGCCATGATGAAGGC 1243

Qy  1201 TCAGAAGCATCTCCTGCTCATCGTCTTGGCCTGAAGCCCAGCCCAGAGGAGCCCTCT 1260
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1244 TCAGAAGCATCTCCTGCTCATCGTCTTGGCCTGAAGCCCAGCCCAGAGGAGCCCTCT 1303

Qy  1261 GCTGCCAGCTTTGAGTACACTATCCTGACCCCCAGTCTTGCGTCCATGGACA 1320
         ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1304 GCTGCCAGCTTTGAGTACACTATCCTGACCCCCAGTCTTGCGTCCATGGACA 1363

Qy  1321 CTGTGCCTGAGCTGCCCCCTACCCCACCCAAAGTACCTGTGTACCTGTGGTATCT 1380
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1364 CTGTGCCCTGAGCTGCCCCCTACCCCACCCAAAGTACCTGTGTACCTTGTGGTATCT 1423

Qy  1381 GACTCTGGCATCTCAACTGACTACAGCTCAGGGGACTCCCAGGGAGCCCAAGGGGGCTTA 1440
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1424 GACTCTGGCATCTCAACTGACTACAGCTCAGGGGACTCCCAGGGAGCCCAAGGGGGCTTA 1483

Qy  1441 TCCGATGGCCCCTACTCCAACCCTTATGAGAACAGCCTTATCCGCCGCTGAGCCTCTG 1500
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1484 TCCGATGGCCCCTACTCCAACCCTTATGAGAACAGCCTTATCCGCCGCTGAGCCTCTG 1543

Qy  1501 CCCCCAGCTATGTGGCTTGCTCTTAG 1527
         ||||||||||||||||||||||||||
Db  1544 CCCCCAGCTATGTGGCTTGCTCTTAG 1570
```

RESULT 2
CR620075

```
LOCUS       CR620075                1797 bp    mRNA    linear   HTC 21-JUL-2004
DEFINITION  full-length cDNA clone CS0DI016YL05 of Placenta Cot 25-normalized
            of Homo sapiens (human).
ACCESSION   CR620075
VERSION     CR620075.1  GI:50500882
KEYWORDS    HTC; CNSLT_cDNA.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 1797)
  AUTHORS   Li,W.B., Gruber,C., Jessee,J. and Polayes,D.
  TITLE     Full-length cDNA libraries and normalization
  JOURNAL   Unpublished
  REMARK    Contact : Feng Liang Email : fliang@lifetech.com URL :
            http://fulllength.invitrogen.com/ Invitrogen Corporation 1600
            Faraday Avenue
REFERENCE   2  (bases 1 to 1797)
  AUTHORS   Genoscope.
  TITLE     Direct Submission
  JOURNAL   Submitted (20-JUL-2004) Genoscope - Centre National de Sequencage :
            BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
            - Web : www.genoscope.cns.fr)
COMMENT     1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
            end enriched, double-strand cDNA was digested with Not I and cloned
            into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library
            was normalized. Library was constructed by Life Technologies, a
            division of Invitrogen.
FEATURES             Location/Qualifiers
     source          1..1797
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="CS0DI016YL05"
                     /tissue_type="Placenta Cot 25-normalized"
                     /plasmid="pCMVSPORT_6"
ORIGIN

Query Match       100.0%;   Score 1527;   DB 3;   Length 1797;
Best Local Similarity   100.0%;   Pred. No. 0;
Matches 1527;   Conservative   0;   Mismatches   0;   Indels   0;   Gaps   0;
```

```
Qy      1 ATGGACCACCTCGGGGCGTCCCCTCTGGCCCCCAAGGTCGGCTCCCCTTGTCTCCTGCTGCT   60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     93 ATGGACCACCTCGGGGCGTCCCCTCTGGCCCCCAAGGTCGGCTCCCCTTGTCTCCTGCTGCT  152

Qy     61 GGGGCCGCCTGGGCCGCCCGGGGCGCCCCGCCTAACCTCCCGGACCCCCAAGTTCGAGAGCAAAGCGGC  120
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    153 GGGGCCGCCTGGGCCGCCCGGGGCGCCCCGCCTAACCTCCCGGACCCCCAAGTTCGAGAGCAAAGCGGC  212

Qy    121 TTGCTGGGGCCCGGGGGCCGAGAGAGCTTCTGCTGCTTCACCGAGCGGTTGGAGGACTTG  180
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    213 TTGCTGGGGCCCGGGGGCCGAGAGAGCTTCTGCTGCTTCACCGAGCGGTTGGAGGACTTG  272

Qy    181 GTGTGTTTCTGGAGGAAGCGGCGAGCGGCTGGGGTGGGCCCGGCAACTACAGCTTCTCC  240
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    273 GTGTGTTTCTGGAGGAAGCGGCGAGCGGCTGGGGTGGGCCCGGCAACTACAGCTTCTCC  332

Qy    241 TACCAGCTCGAGGATGAGCCATGGAGCTGTGTCGCCTGCACCAGGCTCCCACGGCTCGT  300
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    333 TACCAGCTCGAGGATGAGCCATGGAGCTGTGTCGCCTGCACCAGGCTCCCACGGCTCGT  392

Qy    301 GGTGCGGTGCGCGCTTCTGTGTTCGCTGCTACAGCCGACACGTCGCAGCTTCGTGCCCCTA  360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    393 GGTGCGGTGCGCGCTTCTGTGTTCGCTGCTACAGCCGACACGTCGCAGCTTCGTGCCCCTA  452

Qy    361 GAGTAGTGCGGTCACAGCAGCCTCCGGGGCTCCGGCGATATCACCGTGTCATCCACATCAAT  420
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    453 GAGTAGTGCGGTCACAGCAGCCTCCGGGGCTCCGGCGATATCACCGTGTCATCCACATCAAT  512

Qy    421 GAAGTAGTGCTCCTAGACGCCCCGTGGGCGGTTGGCTGACGAGAGCGGC  480
          |||||||||||||||||||||||||||||||||||||||||||||||||
Db    513 GAAGTAGTGCTCCTAGACGCCCCGTGGGCGGTTGGCTGACGAGAGCGGC  572

Qy    481 CACGTAGTGTTGCGCTGGCTGCTGCTCCCGCCTGAGACACCCATGACGTCTCACATCCGCTAC  540
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    573 CACGTAGTGTTGCGCTGGCTGCTGCTCCCGCCTGAGACACCCATGACGTCTCACATCCGCTAC  632

Qy    541 GAGGTGACGACGTCTCGGCCGCCGCAACGGGCGCAGGGAGCCTACAGAGGGTGGAGATCCTGGAG  600
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    633 GAGGTGACGACGTCTCGGCCGCCGCAACGGGCGCAGGGAGCCTACAGAGGGTGGAGATCCTGGAG  692

Qy    601 GGCCGCCACCGAGTGTGTGCTGAGCAACCTGCCGGGGCGCCGGACGCGTACACCTTCGCCGTC  660
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    693 GGCCGCCACCGAGTGTGTGCTGAGCAACCTGCCGGGGCGCCGGACGCGTACACCTTCGCCGTC  752

Qy    661 CGCGCGGCGTATGCGCTGAGCCGAGCTTCGGCGGCTTCGGAGCCTGTGTCGCGAGCCGTG  720
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    753 CGCGCGGCGTATGCGCTGAGCCGAGCTTCGGCGGCTTCGGAGCCTGTGTCGCGAGCCGTG  812

Qy    721 TCGCTGCTGACGCCTAGCGACCTGGACCCCCATCCTGACGCTCTCCTGACCGTCTGCTCCATCCTGTG  780
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    813 TCGCTGCTGACGCCTAGCGACCTGGACCCCCATCCTGACGCTCTCCTGACCGTCTGCTCCATCCTGTG  872

Qy    781 GTCATCCTGGTGCTGCTGACCGTGCTGCTGCTCTCCACCGCGGGGCTCGAAGCAG  840
          ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    873 GTCATCCTGGTGCTGCTGACCGTGCTGCTGCTCTCCACCGCGGGGCTCGAAGCAG  932

Qy    841 AAGATCTGGCCTGGCATCCCGGAGCCCAGAGAGGGAGTTTGAAGGCCTCTTCACCACCAC  900
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    933 AAGATCTGGCCTGGCATCCCGGAGCCCAGAGAGGGAGTTTGAAGGCCTCTTCACCACCAC  992

Qy    901 AAGGGTAACTTCCAGCTGTGGCTGTACCAGAATGATGGCTGCTGCTGTGTGGAGCCCCTGC  960
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    993 AAGGGTAACTTCCAGCTGTGGCTGTACCAGAATGATGGCTGCTGCTGTGTGGAGCCCCTGC 1052

Qy    961 ACCCCCTTCACGGAGGACCCACCTGCTTCCCTGGAAGTCCTCTGCAGACGCGTGCTGGGGG 1020
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1053 ACCCCCTTCACGGAGGACCCACCTGCTTCCCTGGAAGTCCTCTGCAGACGCGTGCTGGGGG 1112

Qy   1021 ACGATGCAGGCAGTGGGACAGATGGAAGCGGCGGGACAGATGATGAGGGCGCCCCTGCTGGGC 1080
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1113 ACGATGCAGGCAGTGGGACAGATGGAAGCGGCGGGACAGATGATGAGGGCGCCCCTGCTGGGC 1172

Qy   1081 AGTGAGCATGCCCAGGATACCTATCTGTGCTGGACAAATGGTGCTGCCCCGGGAACCCG 1140
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1173 AGTGAGCATGCCCAGGATACCTATCTGTGCTGGACAAATGGTGCTGCCCCGGGAACCCG 1232

Qy   1141 CCCAGTGAGGACCTCCCAGGGCCTGGTGCAGTGTGGCAGTAGTGGCATAGTGGCGATGGAAGGC 1200
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1233 CCCAGTGAGGACCTCCCAGGGCCTGGTGCAGTGTGGCAGTAGTGGCATAGTGGCGATGGAAGGC 1292

Qy   1201 TCAGAAGCATCCTCCTGCTCGTTGGCCTCGAAGCCACCAGAGCCCAGAGGAGGGAGCCTCT 1260
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1293 TCAGAAGCATCCTCCTGCTCGTTGGCCTCGAAGCCACCAGAGCCCAGAGGAGGGAGCCTCT 1352

Qy   1261 GCTGCCAGCTTTGAGTACACTACATCCTGACCCACCCCAGCTCCCAGCTCTTGGGTCCATGGACA 1320
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1353 GCTGCCAGCTTTGAGTACACTACATCCTGACCCACCCCAGCTCCCAGCTCTTGGGTCCATGGACA 1412

Qy   1321 CTGTGCCTGAGCTGCGCCCCTACCCCACCCCACCCCAGCTAAAGTACCTGTGTCGTATCT 1380
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1413 CTGTGCCTGAGCTGCGCCCCTACCCCACCCCACCCCAGCTAAAGTACCTTGTGTATCT 1472

Qy   1381 GACTCTGGCATCTCAACTGACTACAGCTCAGGGGGACTCCCAGGGGAGCCCAAGGGGCTTA 1440
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1473 GACTCTGGCATCTCAACTGACTACAGCTCAGGGGGACTCCCAGGGGAGCCCAAGGGGCTTA 1532

Qy   1441 TCCGATGGCCCCTACTTCAAACCCTTATGAGAACAGACCCTTATCCCAGCCGCTGAGCCTCTG 1500
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1533 TCCGATGGCCCCTACTTCAAACCCTTATGAGAACAGACCCTTATCCCAGCCGCTGAGCCTCTG 1592

Qy   1501 CCCCCCAGCTATGTGGCTTGCTCTTAG 1527
          |||||||||||||||||||||||||||
Db   1593 CCCCCCAGCTATGTGGCTTGCTCTTAG 1619
```

RESULT 3
CR618473
LOCUS   CR618473   1811 bp   mRNA   linear   HTC 21-JUL-2004
DEFINITION full-length cDNA clone CS0DM009YC22 of Fetal liver of Homo sapiens
            (human).
ACCESSION   CR618473
VERSION   CR618473.1 GI:50499280
KEYWORDS   HTC; CNSLT cDNA.
SOURCE   Homo sapiens (human)
  ORGANISM Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 1811)
  AUTHORS   Li,W.B., Gruber,C., Jessee,J. and Polayes,D.
  TITLE   Full-length cDNA libraries and normalization
  JOURNAL   Unpublished
  REMARK   Contact : Feng Liang Email : fliang@lifetech.com URL :
            http://fulllength.invitrogen.com/ InVitroGen Corporation 1600
            Faraday Avenue
REFERENCE   2 (bases 1 to 1811)
  AUTHORS   Genoscope.
  TITLE   Direct Submission
  JOURNAL   Submitted (20-JUL-2004) Genoscope - Centre National de Sequencage :
            BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
            - Web : www.genoscope.cns.fr)
COMMENT   1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
            end enriched, double-strand cDNA was digested with Not I and cloned
            into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library
            was normalized. Library was constructed by Life Technologies, a
            division of Invitrogen.
FEATURES   Location/Qualifiers
     source   1..1811
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="CS0DM009YC22"
                     /tissue_type="Fetal liver"
                     /plasmid="pCMVSPORT_6"
ORIGIN

Query Match   100.0%; Score 1527; DB 3; Length 1811;

Best Local Similarity 100.0%; Pred. No. 0;
Matches 1527; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy      1 ATGACCACCTCGGGGGCGTCCCTCTGGCCCCAGTCTGGCTCCCTTGTCTCCTGCTGCT 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     79 ATGACCACCTCGGGGGCGTCCCTCTGGCCCCAGTCTGGCTCCCTTGTCTCCTGCTGCT 138

Qy     61 GGGGCCGCCTGGGGCGTCCCCGGCCCTAACCTCCCGGACCCCAGTTGAGAGCAAAGCGGCC 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    139 GGGGCCGCCTGGGGCGTCCCCGGCCCTAACCTCCCGGACCCCAGTTCGAGAGCAAAGCGGCC 198

Qy    121 TTGCTGGGGCCCGGGGGCCCGAGAGCTTCTGTGCTTCACCGAGCGGTTGGAGGACTTG 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    199 TTGCTGGGGCCCGGGGGCCCGAGAGCTTCTGTGCTTCACCGAGCGGTTGGAGGACTTG 258

Qy    181 GTGTTTCTGGGAGGAAGCGGGAGCGGTGGGTGGGGCCGGGCAACTACAGCTTCTCC 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    259 GTGTTTCTGGGAGGAAGCGGGAGCGGTGGGTGGGGCCGGGCAACTACAGCTTCTCC 318

Qy    241 TACCAGCTCGAGGATGAGCGCTGGTTCTCGCTGCCTACAGCGTCGAGTTCGTGCCTCGT 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    319 TACCAGCTCGAGGATGAGCGCTGGTTCTCGCTGCCTACAGCCGTCGAGTTCGTGCCTCGT 378

Qy    301 GGTGCGGTGCGCTTCTGGTTGTTCGCCTACAGCCGACACGTCGAGTTCGTGTGCCCCTA 360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    379 GGTGCGGTGCGCTTCTGGTTGTTCGCCTACAGCCGACACGTCGAGTTCGTGTGCCCCTA 438

Qy    361 GAGTTGGCGCGTCACAGAGCCTTCGGCGGCTCCGGCGATATCACCGTGTCATCCACATCAAT 420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    439 GAGTTGCGCGTCACAGAGCCTTCGGCGGCTCCGGCGATATCACCGTGTCATCCACATCAAT 498

Qy    421 GAAGTAGTAGTCGCTCCTAGACGCGGGGGGCTTGGTGGGGGCGGTGCGCGGTTGGCTGACGAGAGCGGC 480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    499 GAAGTAGTAGTCGCTCCTAGACGCGGGGGGCTTGGTGGGGGCGGTGCGCGGTTGGCTGACGAGAGCGGC 558

Qy    481 CACGTAGTGTTGCGCTGGCCTGCGCCGCCTGAGACGTCTCACATCCGCTAC 540
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    559 CACGTAGTGTTGCGCTGGCCTGCGCCGCCTGAGACGTCTCACATCCGCTAC 618

Qy    541 GAGGTGACGACCTCTCGGCCGCCGCAACGGGCCAGGGAGCGCTACAGAGGGTGGAGATCTCCTGGAG 600
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    619 GAGGTGGACGCTCTCGGCCGCCGCAACGGGCCAGGGAGCGCTACAGAGGGTGGAGATCTCCTGGAG 678

Qy    601 GGCCGCCACGAGTGTGTGCTGCTGAGCAACTGCGGGGGCCGGACGCGCTACACCTTCGCCGTC 660
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    679 GGCCGCCACGAGTGTGTGCTGCTGAGCAACTGCGGGGGCCGGACGCGCTACACCTTCGCCGTC 738

Qy    661 CGGCGCGTATGGCTGAGCCGAGCTTCGCGGCGTTCGGCGGAGCCTGTG 720
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    739 CGGCGCGTATGGCTGAGCCGAGCTTCGCGGCGTTCGGCGGAGCCTGTG 798

Qy    721 TCGCTGCTGACGGCCTAGCGACCTGGACCCCTCATCCTGACCGCTCTCCCTCATCCTCGTG 780
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    799 TCGTTGCTGACGCGCCTAGCGACCTGGACCCCTCATCCTGACCGCTCTCCCTACATCCTCGTG 858

Qy    781 GTCATCCGGTGCTGCTGACGGGCTGCTGCTGCTCTCCGGGGGCTGCTGAAGCAG 840
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    859 GTCATCCGGTGCTGCTGACCGGGCTGCTGCTGCTCTCCGGGGGCTGCTGAAGCAG 918

Qy    841 AAGATCTGGCCTGCATCCCGAGCGCCAGAGAGCGAGTTTGAAGGCCTCTTCACCACCCAC 900
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    919 AAGATCTGGCCTGCATCCCGAGCGCCAGAGAGCGAGTTTGAAGGCCTCTTCACCACCCAC 978

Qy    901 AAGGGTAACTTCACGGAGACCACCTGCTGTCGTACCAGAATGATGGCTGCCTGGTGGTGACGGCCCCTGC 960
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    979 AAGGGTAACTTCACGGAGGACCACCTGCTGTCGTACCAGAATGATGGCTGCCTGGTGGTGACGGCCCCCTGC 1038

Qy    961 ACCCCTTCACGGAGGACCACCACCTGCTTCCTCGGAAGTCCTCTGAGAGCGGCGTGCTGGGGG 1020
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1039 ACCCCCTTCACGGAGGAGACCACCACCTGCTTCCCTGGAAGTCCTCTGGAAGTCCTCTTCCTCGGAAGTCCTCTGGGG 1098

Qy   1021 ACGATGCAGGCAGTGGAGCGCGGGACAGATGATGAGGGCCCCTGCTGGAGCCAGTGGGC 1080
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1099 ACGATGCAGGCAGTGGAGCGCGGGACAGATGATGAGGGCCCCTGCTGGAGCCAGTGGGC 1158
```

```
Db   1099 ACGATGCAGGCAGTGGAGCGCGGGACAGATGATGAGGGCCCCTGCTGGAGCCAGTGGGC 1158
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   1081 AGTGAGCACATGCCCAGGATACCTATCTGGTGCTGACAAATGGTTGTGCTGCCCGGAACCCG 1140
          |||||||||||||||||||||||||||||||||||||||||||||||||
Db   1159 AGTGAGCACATGCCCAGGATACCTATCTGGTGCTGACAAATGGTTGTGCTGCCCGGAACCCG 1218

Qy   1141 CCCAGTGAGGACCTCCCAGGGCCTGGTGTGGCAGTGTGGACATAGTGGCATGGATGAAGGC 1200
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1219 CCCAGTGAGGACCTCCCAGGGCCTGGTGTGGCAGTGTGGACATAGTGGCCATGGATGAAGGC 1278

Qy   1201 TCAGAAGCATCCTCCTGCTCATCTGCTTTGGCCTCGAAGCCCAGCCCAGAGGGAGCCTCT 1260
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1279 TCAGAAGCATCCTCCTGCTCATCTGCTTTGGCCTCGAAGCCCAGCCCAGAGGGAGCCTCT 1338

Qy   1261 GCTGCCAGTTTGAGTACACTATCCTGGACCCCACCCAGCTCTTGCGTCCATGGACA 1320
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1339 GCTGCCAGTTTGAGTACACTATCCTGGACCCCACCCAGCTCTTCGTCCATGGACA 1398

Qy   1321 CTGTGCCCTGAGCTGCCCCCTACCCCACCCACCTAAAGTACCTGTACCTTGTGGTATCT 1380
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1399 CTGTGCCCTGAGCTGCCCCCTACCCCACCCACCTAAAGTACCTGTACCTTGTGGTATCT 1458

Qy   1381 GACTCTGGCATCTCAACTGACTACAGCTCAGGGGACTCCCAGGGAGGCCCAAGGGGGCTTA 1440
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1459 GACTCTGGCATCTCAACTGACTACAGCTCAGGGGACTCCCAGGGAGGCCCAAGGGGGCTTA 1518

Qy   1441 TCCGATGGCCCCTACTCCAACCCTTATGAGAACAGACCCTTATCCCGAGCCTCTG 1500
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1519 TCCGATGGCCCCTACTCCAACCCTTATGAGAACAGACCCTTATCCCGAGCCTCTG 1578

Qy   1501 CCCCCAGCTATGTGGCTTGCTCTTAG 1527
          ||||||||||||||||||||||||||
Db   1579 CCCCCAGCTATGTGGCTTGCTCTTAG 1605
```

RESULT 4
AY414846

LOCUS   AY414846   1527 bp   DNA   linear   GSS 17-DEC-2003
DEFINITION Homo sapiens EPOR gene, VIRTUAL TRANSCRIPT, partial sequence,
            genomic survey sequence.
ACCESSION   AY414846
VERSION   AY414846.1 GI:39770805
KEYWORDS   GSS.
SOURCE   Homo sapiens (human)
  ORGANISM Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 1527)
  AUTHORS   Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A.,
            Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
            Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
            Adams,M.D. and Cargill,M.
  TITLE   Inferring nonneutral evolution from human-chimp-mouse orthologous
            gene trios
  JOURNAL   Science 302 (5652), 1960-1963 (2003)
   PUBMED   14671302
REFERENCE   2 (bases 1 to 1527)
  AUTHORS   Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A.,
            Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
            Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
            Adams,M.D. and Cargill,M.
  TITLE   Direct Submission
  JOURNAL   Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,
            Rockville, MD 20850, USA
COMMENT   This sequence was made by sequencing genomic exons and ordering
            them based on alignment.
FEATURES   Location/Qualifiers
     source   1..1527
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
     gene   <1..>1527
                     /gene="EPOR"

ORIGIN   /locus_tag="HCMS360"

Query Match   95.3%;   Score 1455;   DB 9;   Length 1527;
Best Local Similarity 95.3%;   Pred. No. 0;
Matches 1455; Conservative 0; Mismatches 72; Indels 0; Gaps 0;

```
Qy     1 ATGGACCACCTCGGGGCGTCCCTCTGGCCCAGGTCGGTCTCCTTTGTCTCCTGCTCGCT   60
         ||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1 ATGGACCACCTCGGGGCGTCCCTCTGGCCCAGGTCGGTCTCCTTTGTCTCCTGCNNNNN   60

Qy    61 GGGGCCGCCTGGGCGCCCCCGCCTAACCTCCCGGACCCCAAGTTCGAGAGCAAAGCGGCC   120
         ||||||||||||||||||||||||||||                        |||||
Db    61 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNCCGGCC   120

Qy   121 TTGCTGGGGGCCCGGGGGCCCGAAGAGCTTCTGTGCTTCACCGAGCGGTTGGAGGACTTG   180
         ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   121 TTGCTGGGGGCCCGGGGGCCCGAAGAGCTTCTGTGCTTCACCGAGCGGTTGGAGGACTTG   180

Qy   181 GTGTGTTTCTGGGAGGAAGCGGCGAGCGCTGGGGCTGGGGCCGAGCAACTACAGCTTCCC   240
         ||||||||||||||||||||||||||||||||||||||||||||||||||
Db   181 GTGTGTTTCTGGGAGGAAGCGGCGAGCGCTGGGGCTGGGGCCGAGCAACTACAGCTTCCC   240

Qy   241 TACCAGCTCGAGGATGAGCCATGAAGTGTGTCGCCTGCACCAGGCTCCCACGGCTCGT   300
         ||||||||||||||||                  |||||||||||||||||||||
Db   241 TACCAGCTCGANNNNNNNNNNNGAAGCTGTGTCGCCTGCACCAGGCTCCCACGGCTCGT   300

Qy   301 GGTGCGGTGCGCTTCTGTTGTTCGCTGCTACAGCCGACACGTCGAGCTTCGAGCTGCCCCTA   360
         |||||||||||||||||||||||||||||||||||||||||||||||||||
Db   301 GGTGCGGTGCGCTTCTGTTGTTCGCTGCTACAGCCGACACGTCGAGCTTCGAGCTGCCCCTA   360

Qy   361 GAGTTGCGGCGTCACAGCAGCCGCCTCCGGCGTGGGCGCCCGATATCACCGTGTCATCCACATCAAT   420
         |||||||||||||||     |||||||||||||||||||||||||||||||||
Db   361 GAGTTGCGGCGTCACAGCAGCCGCCTCCGGCGTGGGCGCCCGATATCACCGTGTCATCCACATCAAT   420

Qy   421 GAAGTAGTGCTCCTAGACGCCCCCCGTGGGCGTGTGGCCGGTTGGCTGACGAGAGCCGGC   480
         ||||||||||||||||||||||||||||||||||||||||||||||||||
Db   421 GAAGTAGTGCTCCTAGACGCCCCCCGTGGGCGTGTGGCCGGTTGGCTGACGAGAGCCGGC   480

Qy   481 CACGTAGTTGTTGCGGTGCTGCTGCCCGCTCCCGCGCCTCGACACCCATGAGACGTCTCACATCCGCTAC   540
         ||||||||||||||||||||||||||||||||||||||||||||||||||
Db   481 CACGTAGTTGTTGCGGTGCTGCTGCCCGCTCCCGCGCCTCGACACCCATGAGACGTCTCACATCCGCTAC   540

Qy   541 GAGGTGGACGTCTCGGCCGCCAACGGCCAGGGAGCCGGGTGGAGATCCTGGAG   600
         ||||||||||||||||||||||||||||||||||||||||||||||||||
Db   541 GAGGTGGACGTCTCGGCCGCCAACGGCCAGGGAGCCGGGTGGAGATCCTGGAG   600

Qy   601 GGCGCCACCGAGTGTGTCGAGCAACCTGCGGGGGCCCGAGCGCGCTACACCTTCGCCGTC   660
         |||||||||||||||||||||||||||||||||||||||||||||||||
Db   601 GGCGCCACCGAGTGTGTCGAGCAACCTGCGGGGGCCCGAGCGCGCTACACCTTCGCCGTC   660

Qy   661 CGCGCGGCTATGGCTGAGCCGGAGCTTCGGCGGCCGGCTTCGGCGAGCCTGTG   720
         ||||||||||||||||||||||||||||||||||||||||||||||
Db   661 CGCGCGGCTATGGCTGAGCCGGAGCTTCGGCGGCCGGCTTCGGCGAGCCTGTG   720

Qy   721 TCGCTGCTGACGCCTAGGGAGACCTGGACCCCCTCATCCCCTCATCCTGACGCTGCTCCGTG   780
         ||||||||||||||||||||||||||||||||||||||||||||||||
Db   721 TCGCTGCTGACGCCTAGGGAGACCTGGACCCCCTCATCCCCTCATCCTGACGCTGCTCCGTG   780

Qy   781 GTCATCCTGGTGCTGCTGACCACCGTGCTGCTCGGCGTGCTCTGAAGCAG   840
         |||||||||||||||||||||||||||||||||||||||||||||||
Db   781 GTCATCCTGGTGCTGCTGACCACCGTGCTGCTCGGCGTGCTCTGAAGCAG   840

Qy   841 AAGATGCTGGCCTGGCATCCCGAGCCCAGGAGCGAGTTGAAGGCCTTCACCACCCAC   900
         |||||||||||||||||||||||||||||||||||||||||||||||
Db   841 AAGATCTGGCCTGGCATCCCGAGCCCAGGAGCGAGTTGAAGGCCTTCACCACCCAC   900

Qy   901 AAGGGTAACTTCCAGCTGTGGCTGTACCAGAATGATGGCTGCCTGTGGTGGAGCCCCTGC   960
         |||||||||||||||||||||||||||||||||||||||||||||||
Db   901 AAGGGTAACTTCCAGCTGTGGCTGTACCAGAATGATGGCTGCCTGTGGTGGAGCCCCTGC   960

Qy   961 ACCCCCTTCACGGAGGACCCACCTGCTTCCCTGGAAGTCCTCTCAGAGCGCGTGCTGGGGG   1020
         |||||||||||||||||||||||||||||||||||||||||||||||
Db   961 ACCCCCTTCACGGAGGACCCACCTGCTTCCCTGGAAGTCCTCTCAGAGCGCGTGCTGGGGG   1020

Qy  1021 ACGATGCAGGCAGGCAGCTGAGCCGGCGGGGACAGATGATGAGGGCCCCTGCTGGAGCCCAGTGGGC   1080
         |||||||||||||||||||||||||||||||||||||||||||||||
Db  1021 ACGATGCAGGCAGGCAGCTGAGCCGGCGGGGACAGATGATGAGGGCCCCTGCTGGAGCCCAGTGGGC   1080

Qy  1081 AGTGAGCAATGCCCAGGATACCTATCTGTGCTGGACAAATGGTTGCTGCCCCGGAACCCG   1140
         ||||||||||||||||||||||||||||||||||||||||||||||||
Db  1081 AGTGAGCAATGCCCAGGATACCTATCTGTGCTGGACAAATGGTTGCTGCCCCGGAACCCG   1140

Qy  1141 CCCAGTGAGGAGACCTCCCAGGCCTGGTGCAGTGTGGCAGTGTGGCCATGGCATGGAAGGC   1200
         |||||||||||||||||||||||||||||||||||||||||||||||
Db  1141 CCCAGTGAGGAGACCTCCCAGGCCTGGTGCAGTGTGGCAGTGTGGCCATGGCATGGAAGGC   1200

Qy  1201 TCAGAAGCATCCTCCTGCTCATCTGCTTTGGCCTGCTTTGGCCTCCTCGAAGGCCAGCCAGAAGGGAGCCTCT   1260
         |||||||||||||||||||||||||||||||||||||||||||||||
Db  1201 TCAGAAGCATCCTCCTGCTCATCTGCTTTGGCCTCCTCGAAGCCCAGCCAGAGGGAGCCCTCT   1260

Qy  1261 GCTGCCAGCTTTGAGTACACTATCCTGGACCCCAGCTCCTTGGGTCCCATGGACA   1320
         |||||||||||||||||||||||||||||||||||||||||||||||
Db  1261 GCTGCCAGCTTTGAGTACACTATCCTGGACCCCAGCTCCTTGGGTCCCATGGACA   1320

Qy  1321 CTGTGCCTGAGTGCGCCCCCTACCCCACCCCCACCCCCACTAAGTAGTACTTGTGGTGTATCT   1380
         |||||||||||||||||||||||||||||||||||||||||||||||
Db  1321 CTGTGCCTGAGTGCGCCCCCTACCCCACCCCCACTAAGTAGTACTTGTGGTGTATCT   1380

Qy  1381 GACTCTGGCATCTCAACTGACTACAGCTCAGGGGACTCCCAGGGGCCCCAGGGGGCTTA   1440
         |||||||||||||||||||||||||||||||||||||||||||||||
Db  1381 GACTCTGGCATCTCAACTGACTACAGCTCAGGGGACTCCCAGGGGCCCCAGGGGGCTTA   1440

Qy  1441 TCCGATGGCCCCTACTCCAACCCTTATGAGAACAGCCTTATCCCAGCCGCTGAGCCTCTG   1500
         |||||||||||||||||||||||||||||||||||||||||||||||
Db  1441 TCCGATGGCCCCTACTCCAACCCTTATGAGAACAGCCTTATCCCAGCCGCTGAGCCTCTG   1500

Qy  1501 CCCCCAGCTATGTGGCTTGCTCTTAG   1527
         ||||||||||||||||||||||
Db  1501 CCCCCAGCTATGTGGCTTGCTCTTAG   1527
```

RESULT 5
CR610088
LOCUS   CR610088   1870 bp   mRNA   linear   HTC 21-JUL-2004
DEFINITION full-length cDNA clone CS0DC003YF02 of Neuroblastoma Cot
           25-normalized of Homo sapiens (human).
ACCESSION CR610088
VERSION   CR610088.1 GI:50490895
KEYWORDS   HTC; CNSLT_cDNA.
SOURCE   Homo sapiens (human)
ORGANISM   Homo sapiens
           Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
           Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 1870)
AUTHORS   Li,W.B., Gruber,C., Jessee,J. and Polayes,D.
TITLE   Full-length cDNA libraries and normalization
JOURNAL   Unpublished
REMARK   Contact : Feng Liang Email : fliang@lifetech.com URL :
           http://fulllength.invitrogen.com/ InvitroGen Corporation 1600
           Faraday Avenue
REFERENCE 2 (bases 1 to 1870)
AUTHORS   Genoscope.
TITLE   Direct Submission
JOURNAL   Submitted (20-JUL-2004) Genoscope - Centre National de Sequencage :
           BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
           - Web : www.genoscope.cns.fr)
COMMENT   1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
           end enriched, double-strand cDNA was digested with Not I and cloned
           into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library
           was normalized. Library was constructed by Life Technologies, a
           division of Invitrogen.
FEATURES   Location/Qualifiers
     source   1..1870
                     /organism="Homo sapiens"
                     /mol_type="mRNA"

```
ORIGIN
                /db_xref="taxon:9606"
                /clone="CS0DC003YF02"
                /tissue_type="Neuroblastoma Cot 25-normalized"
                /plasmid="pCMVSPORT_6"

Query Match        92.5%;   Score 1413;   DB 3;   Length 1870;
Best Local Similarity   93.6%;   Pred. No. 3.6e-311;
Matches 1527;   Conservative   0;   Mismatches   0;   Indels 104;   Gaps   1;

Qy      1 ATGGACCACCTCGGGGCGTCCCTGCCCCCAGTTCGGCTCGGCTCCCTTTGTCTCCTGCTGCT 60
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     44 ATGGACCACCTCGGGGCGTCCCTGCCCCCAGTTCGGCTCGGCTCCCTTTGTCTCCTGCTGCT 103

Qy     61 GGGGCGCCTGGGCGCCCCCCCGCCTAACCTCCGGACCCCAGTTCGAGAGCAAAGCGGCC 120
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    104 GGGGCGCCTGGGCGCCCCCCCGCCTAACCTCCGGACCCCAGTTCGAGAGCAAAGCGGCC 163

Qy    121 TTGCTGGCGGCCCGGGGGCCCGAGAGCTTCGTGCTTCACCGAGCGGTTGGACGACTTG 180
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    164 TTGCTGGCGGCCCGGGGGCCCGAGAGCTTCGTGCTTCACCGAGCGGTTGGACGACTTG 223

Qy    181 GTGTGTTTCTGGGAGGAAGCGGCGAGCGCTGGGGTGGGCCCGGGCAACTACAGCTTCTCC 240
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    224 GTGTGTTTCTGGGAGGAAGCGGCGAGCGCTGGGGTGGGCCCGGGCAACTACAGCTTCTCC 283

Qy    241 TACCAGCTCGAGGATGAGCCATGGAAGCTGTGTCGCCTGCACCAGGCTCCCACGGCTCGT 300
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    284 TACCAGCTCGAGGATGAGCCATGGAAGCTGTGTCGCCTGCACCAGGCTCCCACGGCTCGT 343

Qy    301 GGTGCGGTGCGCTTCGTGTTCGCTGCTACAGCGCGACACGTCGAGCTTCGAGCTTCGTGCCCCTA 360
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    344 GGTGCGGTGCGCTTCGTGTTCGCTGCTACAGCGCGACACGTCGAGCTTCGAGCTTCGTGCCCCTA 403

Qy    361 GAGTTGGCGCGGTCACAGCAGCCTCCGGCGCTCCGGCGCTCCGGCGATATCACCGTGTCATCCACATCAAT 420
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    404 GAGTTGGCGCGGTCACAGCAGCCTCCGGCGCTCCGGCGCTCCGGCGATATCACCGTGTCATCCACATCAAT 463

Qy    421 GAAGTAGTGCTTCCTAGACGCGCCCCGTGGGCGTGGCGCGGTTGGCTGACGAGAGCGGC 480
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    464 GAAGTAGTGCTTCCTAGACGCGCCCCGTGGGCGTGGCGCGGTTGGCTGACGAGAGCGGC 523

Qy    481 CACGTAGTGTTGCGCTGGCTGGCTCCGGCGCGCCTGCGCGCCGCTGAGACACCCATGAGACGTCTCACATCCGCTAC 540
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    524 CACGTAGTGTGTGCGCTGGCTGGCTCCGGCGCGCCTGCGCGCCGCTGAGACGCGTCTCACATCCGCTAC 583

Qy    541 GAGGTGGACGTCTCGGCCCGGCCGGCCAACGGCGCAGGGAGCGTACAGAG---------------- 584
            ||||||||||||||||||||||||||||||||||||||||||||||||||
Db    584 GAGGTGGACGTCTCGGCCCGGCCGGCCAACGGCGCAGGGAGCGTACAGAG---------------- 643

Qy    585 -------------- 584
Db    644 TCCCCTGATTGGCTCAGTTCCACCAGGGCCCCCCCAGTCATCTATTTTTGTCTGCTA 703

Qy    585 -------------- 584
Db    585 --------GGTGGAGATCCTGGACGGGCCGCACCGAGTGTG 616

Qy    704 CGCGTGCCCTCGCCCTGATTCCGCCCCCCAGTGGAGATCCTGGAGGGCCGCACCGAGTGTG 763
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    764 ...

Qy    617 TGCTGAGCAACCTGGGGGGCCGACGCGGACGCGGACGCGGACGCGTCCGCCGTCCGCCGCGTATGGCTG 676
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    764 TGCTGAGCAACCTGGGGGGCCGACGCGGACGCGGACGCGGACGCGTCCGCCGTCCGCCGCGTATGGCTG 823

Qy    677 AGCCGGAGCTTCGGGCGGCGGCCTTGGAGCGCGGC 736
            |||||||||||||||||||||||||||||||||||
Db    824 AGCCGGAGCTTCGGGCGGCCTTCTGGAGCGCGGCTGTGTCGCTGGAGCCTGGTGTCGCTGGAGCCGCCTA 883

Qy    737 GCGACCTGGACCCCCTCATCCTGACGCTCTCCCATCCTTCGGTCATCCTTCGGTCATCCTTGGTCGTGCTGC 796
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    884 GCGACCTGGACCCCCTCATCCTGACGCTCTCCCCTCGGTCATCCTTCGGTCATCCTTGGTCGTGCTGC 943

Qy    797 TGACCGTGCTCGCCTGCTCTTCCCACCGCCCGGGCCTCTGAAGCAGAAGATCTGGCCTGGCA 856
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    944 TGACCGTGCTCGCCTGCTCTTCCCACCGCCCGGGCCTCTGAAGCAGAAGATCTGGCCTGGCA 1003
```

```
Qy    857 TCCCGAGCCCCAGAGAGCGGAGTTTGAAGGGCTTCCTTCACCACCCACAAGGGTAACTTCCAGC 916
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1004 TCCCGAGCCCCAGAGAGCGGAGTTTGAAGGGCTTCCTTCACCACCCACAAGGGTAACTTCCAGC 1063

Qy    917 TGTGGCTGTACCAGAATGATGGCTGCCTGTGGTGAGCCCTGCACCCCTTCACGGAGG 976
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1064 TGTGGCTGTACCAGAATGATGGCTGCCTGTGGTGAGCCCTGCACCCCTTCACGGAGG 1123

Qy    977 ACCCACCTGCTTCCCTGGAAGTCCTCCAGAGCGGCTGCTGGGGACGATGCAGGCAGTGG 1036
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1124 ACCCACCTGCTTCCCTGGAAGTCCTCCAGAGCGGCTGCTGGGGACGATGCAGGCAGTGG 1183

Qy   1037 AGCCGGGACAGATGATGAGGGCCCCTGCTGAGCCAGTGGCAGTGGGCATGCCCAGG 1096
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1184 AGCCGGGACAGATGATGAGGGCCCCTGCTGAGCCAGTGGCAGTGGGCATGCCCAGG 1243

Qy   1097 ATACCTATCTGGTCTGGACAAATGTTGCTGCCCCGGAACCCGCCCAGTGAGGACCTCC 1156
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1244 ATACCTATCTGGTCTGGACAAATGTTGCTGCCCCGGAACCCGCCCAGTGAGGACCTCC 1303

Qy   1157 CAGGGCCTGGTGCAGTGTGGACATAGTGGCCATGGATGAAGGCTCAGAGAGCATCCTCCT 1216
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1304 CAGGGCCTGGTGCAGTGTGGACATAGTGGCCATGGATGAAGGCTCAGAGAGCATCCTCCT 1363

Qy   1217 GCTCATCTGCTTTGGCCTCGAAGCCCAGCCTCCAGCCTCTGCGTCCATGGACACTGTGCCCTGAGCTGC 1276
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1364 GCTCATCTGCTTTGGCCTCGAAGCCCAGCCTCCAGCCTCTGCGTCCATGGACACTGTGCCCTGAGCTGC 1423

Qy   1277 ACACTATCCTGGACCCCAGTCTCCAGCTTCTGCGTCTGGTATCTTGGTATCTTGGTGTACCTGTGACTCTGGTCATCTCAA 1336
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1424 ACACTATCCTGGACCCCAGTCTCCAGCTTCTGCGTCTGGTATCTTGGTATCTTGGTGTACCTGTGACTCTGGTCATCTCAA 1483

Qy   1337 CCCCTACCCCACCCCACCCCACCTAAAGTACCTGTACCTGTGTGGTATCTTGGTATCTTGGTGTACCTGTGACTCTGGTCATCTCAA 1396
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1484 CCCCTACCCCACCCCACCCCACCTAAAGTACCTGTACCTGTGTGGTATCTTGGTATCTTGGTGTACCTGTGACTCTGGTCATCTCAA 1543

Qy   1397 CTGACTACAGCTCAGGGGACTCCCAGGGGAGCCCAAGGGGGCTTATCCGATGGCCCCTACT 1456
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1544 CTGACTACAGCTCAGGGGACTCCCAGGGGAGCCCAAGGGGGCTTATCCGATGGCCCCTACT 1603

Qy   1457 CCAACCCTTATGAGAACAGCAGCTTATCCCGCGCCTGAGCCTGAGCCTCCCCCAGCTATGTGG 1516
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1604 CCAACCCTTATGAGAACAGCAGCTTATCCCGCGCCTGAGCCTGAGCCTCCCCCAGCTATGTGG 1663

Qy   1517 CTTGCTCTTAG 1527
            |||||||||||
Db   1664 CTTGCTCTTAG 1674
```

```
RESULT 6
CR613702
LOCUS        CR613702      1681 bp    mRNA    linear    HTC 21-JUL-2004
DEFINITION   full-length cDNA clone CS0DI076YC13 of Placenta Cot 25-normalized
             of Homo sapiens (human).
ACCESSION    CR613702
VERSION      CR613702.1    GI:50494509
KEYWORDS     HTC; CNSLT cDNA.
SOURCE       Homo sapiens (human)
  ORGANISM   Homo sapiens
             Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
             Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE    1  (bases 1 to 1681)
  AUTHORS    Li,W.B., Gruber,C., Jessee,J. and Polayes,D.
  TITLE      Full-length cDNA libraries and normalization
  JOURNAL    Unpublished
  REMARK     Contact : Feng Liang Email : fliang@lifetech.com URL :
             http://fulllength.invitrogen.com/ InVitrogen Corporation 1600
             Faraday Avenue
REFERENCE    2  (bases 1 to 1681)
  AUTHORS    Genoscope.
  TITLE      Direct Submission
  JOURNAL    Submitted (20-JUL-2004) Genoscope - Centre National de Sequencage :
```

COMMENT   BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
              - Web : www.genoscope.cns.fr)
              1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
              end enriched, double-strand cDNA was digested with Not I and cloned
              into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library
              was normalized. Library was constructed by Life Technologies, a
              division of Invitrogen.

FEATURES   Location/Qualifiers
     source   1..1681
              /organism="Homo sapiens"
              /mol_type="mRNA"
              /db_xref="taxon:9606"
              /clone="CS0DI076YC13"
              /tissue_type="Placenta Cot 25-normalized"
              /plasmid="pCMVSPORT_6"

ORIGIN

Query Match   83.3%; Score 1271.6; DB 3; Length 1681;
Best Local Similarity 99.7%; Pred. No. 5.9e-279;
Matches 1274; Conservative   0; Mismatches   4; Indels   0; Gaps   0;

Qy   245 AGCTCGAGGATGAGCCATGAAGCTGTGTCGCTCCACCAGGCGCTCCCACGGCTCGTGGTG 304
          ||||||||| || |||||||||||||||||||||||||||||||||||||||||||||||
Db   404 ATCGAGGGGATGAGCCATGAAGCTGTGTCGCTCCACCAGGCGCTCCCACGGCTCGTGGTG 463

Qy   305 CGGTGCGCTTCTGGTGTTCGCTGCCTACAGCCGACACGTCGAGCTTCGTGCCCCTAGAGT 364
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   464 CGGTGCGCTTCTGGTGTTCGCTGCCTACAGCCGACACGTCGAGCTTCGTGCCCCTAGAGT 523

Qy   365 TGCGGGTCACAGCAGCCTCCGGCGCTCCGGATATCACCGTGTCATCCACTCAATGAAG 424
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   524 TGCGGGTCACAGCAGCCTCCGGCGCTCCGGATATCACCGTGTCATCCACTCAATGAAG 583

Qy   425 TAGTGCTTCCTAGACGCCCCCGTGGGCGTTGGCTGGCTGTGACGAGAGCGGCCACG 484
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   584 TAGTGCTTCCTAGACGCCCCCGTGGGCGTTGGCTGGCTGTGACGAGAGCGGCCACG 643

Qy   485 TAGTGTTGGCGTTGCGCGCTCCGGCCGCAACCGCAGGAGGGTGAGATCTCACATCGCTCACATCCGCTACGAGG 544
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   644 TAGTGTTGGCGTTGCGCGCTCCGGCCGCAACCGCAGGAGGGTGAGATCTCACATCGCTCACATCCGCTACGAGG 703

Qy   545 TGGACGTCTCGGCCGCAACGGCAGGAGGGTGTACAGAGGGCTCCAGGGAGCCCAGGGGGCC 604
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   704 TGGACGTCTCGGCCGCAACGGCAGGAGGGTGTACAGAGGGCTCCAGGGAGCCCAGGGGGCC 763

Qy   605 GCACCGAGTGTGTGCTGAGCAACCTGCGGGGGCCGACCTTCGCCGGTCCGCG 664
          |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   764 GCACCGAGTGTGTGCTGAGCAACCTGCGGGGGCCGACCTTCGCCGGTCCGCG 823

Qy   665 CGGTATGGCTGAGCCGAGCTTCGGCGGTCTTCGGAGCCTGGTCGGAGCCTGGTGTCGC 724
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   824 CGGGTATGGCTGAGCCGAGCTTCGGCGGTCTTCGGAGCGCTGGTCGGAGCCTGGTGTCGC 883

Qy   725 TGCTGACGGCCTAGGCAGTGACCTGAGCGTCTCCCTCATCCTCGTGGTCA 784
          ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   884 TGCTGACGGCCTAGGCAGTGACCTGAGCGTCTCCCTCATCCTCGTGGTCA 943

Qy   785 TCCTGGTGCTGCTGACCGTGCTCGCGCTCCTCCCACCGCGGCTCTGAAGCAGAAGA 844
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   944 TCCTGGTGCTGCTGACCGTGCTCGCGCTCCTCCCACCGCGGCTCTGAAGCAGAAGA 1003

Qy   845 TCTGGCCTGGCATCCGAGCCCAGAGAGCGAGTTTGAAGGCCTCTTCACCACCCACAAGG 904
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1004 TCTGGCCTGGCATCCGAGCCCAGAGAGCGAGTTTGAAGGCCTCTTCACCACCCACAAGG 1063

Qy   905 GTAACTTCCAGCTGCGTGGCGTGTACCAGAATGATGGCTGCCGTGGTGGAGGCCCTGCACCC 964
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1064 GTAACTTCCAGCTGCGTGGCGTGTACCAGAATGATGGCTGCCGTGGTGGAGGCCCTGCACCC 1123

Qy   965 CCTTCACGGAGGAGACCCACCTGCTTCCCTGAAGTCCTTCACAGAGTCCTTCAGAGCGCGCTGCTGGGGGACGA 1024
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1124 CCTTCACGAGGAGACCCACCTGCTTCCCTGAAGTCCTTCACAGAGTCCTTCAGAGCGCGCTGCTGGGGGACGA 1183

Qy   1025 TGCAGGCAGTGGAGCGCGGGGACAGGATGATGAGGGCCCCTGCTGGCAGCCCCTGCTGGCCAGTG 1084
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1184 TGCAGGCAGTGGAGCGCGGGGACAGGATGATGAGGGCCCCTGCTGGGAGCCCAGTGGCCAGTG 1243

Qy   1085 AGCATGCCCAGGATACCTATCTGGTGCTGGACAAATGGTTGCTGCCCGGAACCCGCCCA 1144
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1244 AGCATGCCCAGGATACCTATCTGGTGCTGGACAAATGGTTGCTGCCCGGAACCCGCCCA 1303

Qy   1145 GTGAGGACCTCCAGGCCTGCTGGCAGTGTGGACATAGTGCCATGATGAAGGCTCAG 1204
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1304 GTGAGGACCTCCAGGCCTGCTGGCAGTGTGGACATAGTGCCATGATGAAGGCTCAG 1363

Qy   1205 AAGCATCCTCCTGCTCATCTGCTTTGGCCTCGAAGCCCAGCCCAGAGGGAGGGGAGCCTCTGCTG 1264
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1364 AAGCATCCTCCTGCTCATCTGCTTTGGCCTCGAAGCCCAGCCCAGAGGGAGGGGAGCCTCTGCTG 1423

Qy   1265 CCAGCTTTGAGTACACTATCCTGGACCCCAGCTCCCAGCTCTTGCGTCCATGACACTGT 1324
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1424 CCAGCTTTGAGTACACTATCCTGGACCCCAGCTCCCAGCTCTTGCGTCCATGGACACTGT 1483

Qy   1325 GCCCTGAGCTGCCCCCTACCCCCACCCCACCCACCCACCTAAAGTACCTGTACCTTGTGGTATCTGACT 1384
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1484 GCCCTGAGCTGCCCCCTACCCCCACCCCACCCACCTAAAGTACCTGTACCTTGTGGTATCTGACT 1543

Qy   1385 CTGGCATCCAACTGACTACAGCTCAGGGAGCTCCCAGGGAGCCCAAGGGGGGCTTATCCG 1444
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1544 CTGGCATCCAACTGACTACAGCTCAGGGAGCTCCCAGGGAGCCCAAGGGGGGCTTATCCG 1603

Qy   1445 ATGGCCCTACTCCAACCCTTATGAGAACAGCCTTATCCCAGCCGCTCAGCCTCTGCCCC 1504
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1604 ATGGCCCCTACTCCAACCCTTATGAGAACAGCCTTATCCCAGCCGCTCAGCCTCTGCCCC 1663

Qy   1505 CCAGCTATGTGGCTTGCT 1522
          ||||||||||||||||||
Db   1664 CCAGCTATGTGGCTTGCT 1681

RESULT 7
AY414848
LOCUS   AY414848   1452 bp   DNA   linear   GSS 17-DEC-2003
DEFINITION Mus musculus EPOR gene, VIRTUAL TRANSCRIPT, partial sequence,
            genomic survey sequence.
ACCESSION   AY414848
VERSION   AY414848.1 GI:39770807
KEYWORDS   GSS.
SOURCE   Mus musculus (house mouse)
  ORGANISM Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1 (bases 1 to 1452)
  AUTHORS   Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A.,
            Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
            Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
            Adams,M.D. and Cargill,M.
  TITLE   Inferring nonneutral evolution from human-chimp-mouse orthologous
            gene trios
  JOURNAL   Science 302 (5652), 1960-1963 (2003)
   PUBMED   14671302
REFERENCE   2 (bases 1 to 1452)
  AUTHORS   Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A.,
            Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
            Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
            Adams,M.D. and Cargill,M.
  TITLE   Direct Submission
  JOURNAL   Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,
            Rockville, MD 20850, USA
COMMENT   This sequence was made by sequencing genomic exons and ordering
            them based on alignment.
FEATURES   Location/Qualifiers
     source   1..1452
                     /organism="Mus musculus"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:10090"

```
gene            <1..>1452
                /gene="EPOR"
                /locus_tag="HCM5360"
ORIGIN

Query Match     65.9%;  Score 1005.8;  DB 9;  Length 1452;
Best Local Similarity   81.5%;  Pred. No. 2.2e-218;
Matches 1195;  Conservative  0;  Mismatches 257;  Indels  15;  Gaps  2;

Qy   61 GGGGCCCCTGGGCGCCCCGCTCAACCTCCCGGACCCCCAAGTTCGAGAGCAAAGCGGGC  120
        |||||  |||||| ||||||   |||||||||||||||||||||||||||||||||||||
Db    1 GGGGCAGCCTGGGCACCTTCACCCAGCCTCCCGACCCCCAAGTTCACCGAGGGTTGAGGACTTG  60

Qy  121 TTGCTGCGGCGCCCGGGGCCCGAAGAGCTTCTGTGCTTCACCGAGCGGTTCGGGAGGACTTG  180
        |||||||||||  ||||||||| ||||||||||||||||||||||| |||||||||
Db   61 CTGCTGCGCATCCCGGGGGCTCCGAGGAACTTCTGTGCTTCACCCAACCGTTGAAGAGACTTG  120

Qy  181 GTGTGTTTCTGGGAGGAAGCGGCGAGCGGTGGCGTCCGGATGAC---TTCAACTACAGCTTCTCC  240
        ||||||||||||||||||||||||| |||||   ||||||||| |  ||||||||||||||||
Db  121 GTGTGTTTCTGGGAGGAAGCGGCGGCGTCCGGATGAC---TTCAACTACAGCTTCTCA  177

Qy  241 TACCAGCTCGAGGATGAGCCATGGAAGCTGTCTGCCTGCACCAGGCTCCCACGGCTCGT  300
        ||||||||||||||||||||||  |||| | | ||| || ||||||||||||||||
Db  178 TACCAGCTCGAGGGTGAGTCACGAAAGTCATGTAGCCTGCGACCAGGCTCCCACCGCTCCGC  237

Qy  301 GGTGCCGGTTGCCTTCGCTGGTGTTGCGTCCTACAGCCGACACGTCGAGCTTCGTGCCCCTA  360
        |||||||| ||  ||| || | |||||| ||||||||| ||||||||||| |||||||||
Db  238 GGCTCCCGTGCGTTTCTGGTGTTCACTGCCAACAGCGGACACATCGAGTTTTGTTGTGCCGCTG  297

Qy  361 GAGTTGGCGGTCACAGCAGCCTTCGGCGCGTCCGGGCTTCGCGGATATCACCGTGTCATCCACATCAAT-----  420
        ||||  | |||||  |||| ||||   ||||||||||||  ||   ||||||| |||||||||
Db  298 GAGCTGCGAGGTGACGGAGGGCGTCCGGTCTCCGTCATCACCGCTCATCACCGCATCATCCATA-----  352

Qy  421 GAAGTAGTGCTCCTAGACGCGCCCGTGGGGCTGGTGCGGGCGGGTTGGCTGACGAGAGCGGC  480
              |  |||| ||||||||| |||||||||||||| |||||| ||||||| |||||||||
Db  353 ------TGCTCCTGGACGCGCCCGTGGGGCTGGTGCGGGCGCTGGCGCCGCAGAGGAGGGCAGC  405

Qy  481 CACGTAGTGTTGCGCTGGCTCCCGCGCCGCTGAGACACCCATGACGTCCACATCCGCTAC  540
        |||||||||| ||||||||||||||||||||||||||  |||| |  ||||||||| |||
Db  406 CACGTGGTGCTGCGCTGGCTGCCACCTCCTGGAGCACCCATGATGACCACCCACATCCGATAT  465

Qy  541 GAGGTGACGTCTCGGCGGCGGCAACGCGGCAGGGAGGTAGAGAGGTGGAGATCCTCGGAG  600
        |||||||||||||||||||||||||||||| |||||||| |||||| |||||||| |||
Db  466 GAAGTGGACGTGTCGGCAGGCAACCCGGGCACGAGGAGGACACAAAGGGTGGAGTCCTGGAA  525

Qy  601 GGCGCCACGAGTGTGTGCTGAGCAACCTGCGGGGCGCCGACACGCGGCTACACCTTCGCGTC  660
        |||| ||||||||||||||||||||||||||||||| |||||| |||||||||||||||||
Db  526 GGCCCACTGAGTGTGTTCTGAGCAACCTGCGGGGACGCGCGTACACCTTCGCGTGTT  585

Qy  661 CGGCGGCGTATGGGCTAGCCGGCTTCGGCGGCTTCGGCGCCGGTCGGTGTGGCGGCGCTGTG  720
        ||||| ||||||||||  |||||||||||||||||||| |||||||| |||||| ||||||
Db  586 CGAGCGCGCATGGGCCGCCGGCCGCCAGATTCTCAGCGGCTTCAGCCGAGTCTGAGCCGCCGCG  645

Qy  721 TCGCTGCTGACGACGTCCTAGGCGACCTGGACCCCCTCATCCTGACGCGTGCTCCCCATCCTCGTG  780
        |||||||||||| ||| ||||||||| |||||| |||| |||| ||| |||| |||  ||| |
Db  646 TCAACTACTGACGCGTCTAGCCGAGCCTGGACCCCTCTCCTCCATCTTGACGCGTGTCTCATTCTGGTC  705

Qy  781 GTCATCCGTGGCTGCTGCTGACCCGTGCTGCTGACGCGTGCTCTCCACCCGCGGGGCTCGAAGCAG  840
        |||||||| |||||||||||||||| |||||||||  |||| ||||| ||||||||||||||||
Db  706 CTCATCTCGCTGTTGCTGACGGTTCGGACGGTTCGGACGTTCCCACCGCCGGACTCTGCAGCAG  765

Qy  841 AAGATCTGGCCTGGCATCCCGAGCCCCAGAGAGCGAGTTTGAAGGCCTTCTTCACCACCCAC  900
        ||||||||||||||||| ||| ||||| ||| ||||||||||||  | || ||||||||||
Db  766 AAGATCTGGCCTGGCATCCCAAGCCCCAGAGAGCGAGTTTGAGGGGTCTTTCACCACCCAC  825

Qy  901 AAGGGTAACTTCACGCTGTGCTGTACCAGAATGATGGCTGCCTGGTGAGCCCCTGC  960
        |||||||||||||| ||||| ||||||  ||  | |||  |||||| || |||||
Db  826 AAGGGTAACTTCACGCTGTGCTGCAGCGCTGATGTGTCTGTTCTCTGGTGGTGGAGCCCCGGGC  885

Qy  961 ACCCCCTTCACGGAGGACCCACCTGCTTCCTGCGGAAGTCTTCTCAGAAGCGCTGCTGGGGG  1020
        ||||||  |||||||||| | |||||||||| ||  | |||||||||||||||| ||||||
Db  886 AGCTCCTTCCCCTGAGGATCACCTGCCACCTAGAGGTTCCTCTCAGAAGCCACGCACGCTGGGCA  945
```

```
Qy 1021 ACGATGCGAGGCAGTGGAGCCGGGGACAGATGATGAGGGCCCCCTGCTTGGAGGCCAGTGGGC  1080
        ||||  |||||||||||||||||| |||||||||||||||||||||||||||||||||||||
Db  946 GTGACTCAGGCTGGGGACCCCAGGGCAGATGATGAGGGGCCCTTACTGAGCCGGGCGGTGGGC  1005

Qy 1081 AGTGAGCAGCCAGGATACCTATCTGGTGTCTGGACAAATGGTTGCTGCCCGGAACCCG  1140
        ||||||||||||||||||||||||||||| ||||||||||||||| ||||  ||||||
Db 1006 AGTGAGCAGCCAGGAGACCCAGCACCTACTTCGGTATTGGATAAGTGGTTGCTGCCCCGGACCCCA  1065

Qy 1141 CCCAGTGAGGACCTGCCCCAGGGCCTGGTGGCGAGTGTGGCAGCATAGTGGCGCATAGTGGATGAAGGC  1200
        |||| ||||||| || ||||||||||||| |||||||||||||| ||| || |||||| |||||||
Db 1066 TGCAGTGAGAACCTCTCAGGGGCCTGGGGCCAGTGTGGACCCTGTGACTATGGATGAAGCT  1125

Qy 1201 TCAGAAGCATCCTCCTGCTTCTGGACACCTCCCAGCTTCCGAGCTCCCAGCTCTTGCGTCATGGACA  1260
        |||| | || |||| |||||||| ||||||||||| |||  |||||||||| ||||||||||||
Db 1126 TCAGAAACATCTTCCTGCCCCGTTCTGACTTGGCCTCAAAGCCCAGGCCAGGGCACCTCA  1185

Qy 1261 GCTGCCAGCTTGAGTACACTATCCTGGACCCCAGCTCTCCAGCTCTTCGCGTCTCCATGGACA  1320
        |||||||||||||||| || || |||||||||| ||||| ||| |||| |||| ||| ||||
Db 1186 CCTTCCAGCTTTGAGTACACCACCATCCTGGACCCCCAGCTCTCTCAGCTCCTCGTGCCCTCGGGCA  1245

Qy 1321 CTGTGCCCTGAGCTGCCCCCTCACCTCTGTGCCTCTGCCTCCAGGCTCTCCACGGCTCGT  1380
        ||||||||| ||||| ||  |||||||  | |||||| || ||||||||| |||  |||
Db 1246 CTGCCTCCCGGAGCTACCTCCACTCTCACTCGTGAAGTACCTATACCTTGTGGTGTCC  1305

Qy 1381 GACTCTGGCATCTCAACTGACTACAGCTCAGCTGACTGGGGGCTCAGGAGCCCAGGGCTTA  1440
        ||| || |||||||| |||| || || |  | |||| |||||| ||||||| |||||||||
Db 1306 GATTCTGGCATCCAACAGATTCAACAGTTCGGGGCTCGGGCGGGTGCTCAGGGGGGACTCA  1365

Qy 1441 TCCGATGGCCCCTACTCCAACCTTATGAGAACAGCTTATCCCAGCCGCTGAGCCCTG  1500
        || ||||||||| || |||| |||||||||||  ||||| ||| || |||||||||
Db 1366 TCTGATGGCCCCTACTCCCACCCCTATGAGAACAGCGCTTGTCCCAGACTCAGAGCCTCTG  1425

Qy 1501 CCCCCAGCTATGTGGCTTGCTCTTAG  1527
        |||||  ||||||||| |||||||||
Db 1426 CATCCCGGCTATGTGGCCTGCTCCTAG  1452
```

```
RESULT 8
BX382170/c
LOCUS       BX382170               1082 bp    mRNA    linear   EST 26-APR-2004
DEFINITION  BX382170 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens cDNA
            clone CS0DI076YC13 3-PRIME, mRNA sequence.
ACCESSION   BX382170
VERSION     BX382170.2  GI:46572612
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
ORGANISM    Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 1082)
AUTHORS     Li,W.B., Gruber,C., Jessee,J. and Polayes,D.
TITLE       Full-length cDNA libraries and normalization
JOURNAL     Unpublished (2001)
COMMENT     On May 8, 2003 this sequence version replaced gi:30449031.
            Contact: Genoscope
            Genoscope - Centre National de Sequencage
            2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE
            Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
            1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
            end enriched, double-strand cDNA was digested with Not I and cloned
            into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library
            was normalized. Library was constructed by Life Technologies, a
            division of Invitrogen. This sequence belongs to sequence cluster
            9443.r.
            For more information about this cluster, see
            http://www.genoscope.cns.fr/cdna?s=CS0DI076AB07NP1&c=9443.r.
FEATURES             Location/Qualifiers
     source          1..1082
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="CS0DI076YC13"
                     /tissue_type="PLACENTA COT 25-NORMALIZED"
```

```
Db    161  CCTGTACCTTGTGGTATCTGACTCTGGCATCTCAACTGACTACAGCTCAGGGGACTCCCA  102
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   1422  GGGAGCCCAAGGGGCTTATCCGATGGCCCTCATCCGGCCCCTTATGAGAACAGCCTTAT  1481
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    101  GGGAGCCCAAGGGGCTTATCCGATGGCCCTCATCCGGCCCCTTATGAGAACAGCCTTAT   42

Qy   1482  CCCAGCCGCTGAGCCTCTGCCCCCCAGCTATGTGGCTTGCT  1522
           |||||||||||||||||||||||||||||||||||||||||
Db     41  CCCAGCCGCTGAGCCTCTGCCCCCCAGCTATGTGGCTTGCT     1
```

RESULT 9
AK010968
LOCUS   AK010968   1524 bp   mRNA   linear   HTC 03-APR-2004
DEFINITION   Mus musculus 13 days embryo liver cDNA, RIKEN full-length enriched
               library, clone:2510015H03 product:erythropoietin receptor, full
               insert sequence.
ACCESSION   AK010968
VERSION   AK010968.1 GI:12846772
KEYWORDS   HTC; CAP trapper.
SOURCE   Mus musculus (house mouse)
  ORGANISM   Mus musculus
               Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
               Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1
  AUTHORS   Carninci,P. and Hayashizaki,Y.
  TITLE   High-efficiency full-length cDNA cloning
  JOURNAL   Meth. Enzymol. 303, 19-44 (1999)
  MEDLINE   99279253
   PUBMED   10349636
REFERENCE   2
  AUTHORS   Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K.,
               Itoh,M., Konno,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.
  TITLE   Normalization and subtraction of cap-trapper-selected cDNAs to
               prepare full-length cDNA libraries for rapid discovery of new genes
  JOURNAL   Genome Res. 10 (10), 1617-1630 (2000)
  MEDLINE   20499374
   PUBMED   11042159
REFERENCE   3
  AUTHORS   Shibata,K., Itoh,M., Aizawa,K., Nagaoka,S., Sasaki,N., Carninci,P.,
               Konno,H., Akiyama,J., Nishi,K., Kitsunai,T., Tashiro,H., Itoh,M.,
               Sumi,N., Ishii,Y., Nakamura,S., Hazana,M., Nishine,T., Harada,A.,
               Yamamoto,R., Matsumoto,H., Sakaguchi,S., Ikegami,T., Kashiwagi,K.,
               Fujiwake,S., Inoue,K., Togawa,Y., Izawa,K., Ohara,E., Watahiki,M.,
               Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T., Matsuura,S., Kawai,J.,
               Okazaki,Y., Muramatsu,M., Inoue,Y., Kira,A. and Hayashizaki,Y.
  TITLE   RIKEN integrated sequence analysis (RISA) system--384-format
               sequencing pipeline with 384 multicapillary sequencer
  JOURNAL   Genome Res. 10 (11), 1757-1771 (2000)
  MEDLINE   20530913
   PUBMED   11076861
REFERENCE   4
  AUTHORS   The RIKEN Genome Exploration Research Group Phase II Team and the
               FANTOM Consortium.
  TITLE   Functional annotation of a full-length mouse cDNA collection
  JOURNAL   Nature 409, 685-690 (2001)
REFERENCE   5
  AUTHORS   The FANTOM Consortium and the RIKEN Genome Exploration Research
               Group Phase I & II Team.
  TITLE   Analysis of the mouse transcriptome based on functional annotation
               of 60,770 full-length cDNAs
  JOURNAL   Nature 420, 563-573 (2002)
REFERENCE   6 (bases 1 to 1524)
  AUTHORS   Adachi,J., Aizawa,K., Akahira,S., Akimura,T., Arai,A., Aono,H.,
               Arakawa,T., Bono,H., Carninci,P., Fukuda,S., Fukunishi,Y.,
               Furuno,M., Hanagaki,T., Hara,A., Hayatsu,N., Hiramoto,K.,
               Hiraoka,T., Hori,F., Imotani,K., Ishii,Y., Itoh,M., Izawa,M.,
               Kasukawa,T., Kato,H., Kawai,J., Kojima,Y., Konno,H., Kouda,M.,
               Koya,S., Kurihara,C., Matsuyama,T., Miyazaki,A., Nishi,K.,
               Nomura,K., Numazaki,R., Ohno,M., Okazaki,Y., Okido,T., Owa,C.,
               Saito,H., Saito,R., Sakai,C., Sakai,K., Sano,H., Sasaki,D.,
               Shibata,K., Shibata,Y., Shinagawa,A., Shiraki,T., Sogabe,Y.,
               Suzuki,H., Tagami,M., Tagawa,A., Takahashi,F., Tanaka,T.,
```

```
               /clone_lib="Homo sapiens PLACENTA COT 25-NORMALIZED"
               /note="1st strand cDNA was primed with a NotI-oligo(dT)
               primer. Five prime end enriched, double-strand cDNA was
               digested with Not I and cloned into the Not I and EcoR V
ORIGIN         sites of the pCMVSPORT 6 vector. Library was normalized."

Query Match       64.0%;   Score 977;   DB 5;   Length 1082;
Best Local Similarity   97.5%;   Pred. No. 7.6e-212;
Matches 1034;   Conservative 10;   Mismatches 10;   Indels 7;   Gaps 5;

Qy    462  GTTGGTGACGAGAGACGGGCCACGTAGTGTTGCGCTGGCT  521
           ||||||||||||:|||||||: |||||||: ||||:||||:
Db   1054  GTTGGYTGACAGAGAGS-GCCACGTAGTGTGSGGYTG--TCCCGCCGCCTGAAAMACCCAT  998

Qy    522  GACGTCTCACATCCGCTACGAGGTGACGTCTCGGCGGCAACGGCGCAGGGAGGAGCGTACA  581
           :::||||||||||||||| :| || | ||||||||||||||||||||||||||||||||||
Db    997  -RMSTCTCACATCCGCTACGA-GTGACGTCTCGGCGGCAACGGCGCAGGGAGGAGCGTACA  940

Qy    582  GAGGGTGGAGATCCTGGAGGCCGCCGACCGAGTGTGTCTGAGCAACCTGCGGGGCCGGAC  641
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    939  GAGGGTGGAGATCCTGGAGGCCGCCGACCGAGTGTGTCTGAGCAAACTGCGGGGCCGGAC  880

Qy    642  GCGCTACACCTTCGCCGCGTCCGCGCGGGTATGGCTGAGCCGAGCTTCGGCGGCTTCTGGAG  701
           ||||||||||||||||||| |||||||||||||||||||||| ||||||||||||||||||
Db    879  GCGCTACACCTTCGCCGCGTCCGCGCGGGTATGGCTGAGCCGAGCTTCGGCGGCTTCTGGAG  820

Qy    702  CGCCTGGTCGGAGCCTGTGTCGCTGCTAGCGCTAGCGCCTGGACCCCCTCATCCTGAC  761
           ||||||||||||||:| ||||||||||||:||||||||||||||||||||||||||||
Db    819  CGCCTGGTCGGAGCCTGTGTCGYCGTGACGCCTAGCGACCTGGACCCCCTCATCCTGAC  760

Qy    762  GCTCCCCTCATCCTCGTGTCGTGGTCATCCTCGTGGTGTCGCTGCTCGCGCTGCTCTCCCA  821
           ||||||||||||||||| ||||||||||||||||||| |||||||||||||||||||||||
Db    759  GCTCCCCTCATCCTCGTGTCGTGGTCATCCTGGTGTCGCTGCTGTCGCTGCTGT--GGTGTCGTCTCCCA  702

Qy    822  CCGCCGGGCTCTGAAGCAGAAGATCTGGCCTGGACATCCCGAGCCCAGAGAGCGAGTTTGA  881
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    701  CCGCCGGGCTCTGAAGCAGAAGATCTGGCCTGGACATCCCGAGCCCAGAGAGCGAGTTTGA  642

Qy    882  AGGCCTCTTCACCACCCACAAGGTAACTTCCAGCTGTGCTGCTACCAGAATGATGGCTG  941
           ||||||||||||||||||||||||| |||||||||||||||||||||||||||||||||
Db    641  AGGCCTTCACCACCCACCACAGGGTAACTTCCAGCTGTGCTGCTGTTACCAGAATGATGGCTG  582

Qy    942  CCTGTGGTGGAGCCCCGGAGACCCACCTGCTTCCCTGGAAGTCCT  1001
           ||||||||||||||||||||||||||||||||||||||||||||||
Db    581  CCTGTGGTGGAGCCCCCGGAGACCCACCTGCTTCCCTGGAAGTCCT  522

Qy   1002  CTCAGAGCCGCTGCTGGGGGACGATGCAGGCAGTGAGCCCGGGACAGATGAGGGCCC  1061
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    521  CTCAGAGCCGCTGCTGGGGGACGATGCAGGCAGTGAGCCCGGGACAGATGAGGGCCC  462

Qy   1062  CCTGCTGGAGCCAGTGGCCAGTGAGCCAGGATCGAGCGCAGCATGCCCCAGGATACCTGGTGCTGGACAAATG  1121
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    461  CCTGCTGGAGCCAGTGGCCAGTGAGCCAGGATCGAGCATGCCCCAGGATACCTGGTGCTGGACAAATG  402

Qy   1122  GTTGCTGCCCCGGAACCCGGACCTCCCCAGTGGCCTGCTGGTGCAGTGTGGACAT  1181
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    401  GTTGCTGCCCCCGGAACCCCGCCAGTGGCCTGCTGGTGCAGTGTGGACAT  342

Qy   1182  AGTGGCCATGGATGAAGGCTCAGAGAGCATCTCTCTGCTTTCGCCTCGAAGCC  1241
           ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    341  AGTGGCCATGGATGAAGGCTCAGAGAGCATCTCTCTGCTTTCGCCTCGAAGCC  282

Qy   1242  CAGCCCAGAGGGAGCAGCCTCGTGCTGCAGCTTTGAGTACACTACTATCCTGGGACCCCAGCTCCCA  1301
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    281  CAGCCCAGAGGGGAGCCTCGTGCTGCAGCTTTGAGTACACTACTATCCTGGGACCCCAGCTCCCA  222

Qy   1302  GCTCTTGGGTCCATGGACACTGCCCTGAGCTGCCCCTCACCCCCCTCACCCTAAAGTA  1361
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    221  GCTCTTGGGTCCATGGACACTGCCCTGAGCTGCCCCTCACCCCCCTCACCCTAAAGTA  162

Qy   1362  CCTGTACCTTGTGGTATCTGACTCTGGCATCTCAACTGACTACAGCTCAGGGGACTCCCA  1421
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

TITLE   Tejima,Y., Toya,T., Yamamura,T., Yasunishi,A., Yoshida,K., Yoshino,M., Muramatsu,M. and Hayashizaki,Y.

JOURNAL   Direct Submission
Submitted (10-JUL-2000) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute: 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail:genome-res@gsc.riken.jp, URL:http://genome.gsc.riken.jp/, Tel:81-45-503-9222, Fax:81-45-503-9216)

COMMENT   Please visit our web site (http://genome.gsc.riken.jp/) for further details.
cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN.
Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. First strand cDNA was primed with a primer [5' GAGAGAGAGAGGATCCAAGAGCTCTTTTTTTTTTTTTVN 3'], cDNA was prepared by using trehalose thermo-activated reverse transcriptase. and subsequently enriched for full-length by cap-trapper. Second strand cDNA was prepared with the primer adapter of sequence [5' GAGAGAGAGATTCTCGAGTTAATTAAATTAATCCCCCCCCCCCC 3']. cDNA was cleaved with XhoI and SstI. Cloning sites, 5' end: XhoI; 3' end: SstI. Host: SOLR.

FEATURES   Location/Qualifiers
source   1..1524
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="FANTOM DB:2510015H03"
/db_xref="taxon:10090"
/clone="2510015H03"
/tissue_type="liver"
/clone_lib="RIKEN full-length enriched mouse cDNA library"
/dev_stage="13 days embryo"
misc_feature   <1..1377
/note="erythropoietin receptor (MGD|MGI:95408, GB|NM_010149, evidence: BLASTN, 99%, match=1523)
putative"

ORIGIN

Query Match   62.2%;   Score 949.6;   DB 3;   Length 1524;
Best Local Similarity 82.0%;   Pred. No. 1.4e-205;
Matches 1131;   Conservative 0;   Mismatches 244;   Indels 5;   Gaps 3;

```
SOURCE       Homo sapiens (human)
ORGANISM     Homo sapiens
             Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
             Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE    1 (bases 1 to 1320)
AUTHORS      Li,W.B., Gruber,C., Jessee,J. and Polayes,D.
TITLE        Full-length cDNA libraries and normalization
JOURNAL      Unpublished
REMARK       Contact : Feng Liang Email : fliang@lifetech.com URL :
             http://fulllength.invitrogen.com/ InvitroGen Corporation 1600
             Faraday Avenue
REFERENCE    2 (bases 1 to 1320)
AUTHORS      Genoscope.
TITLE        Direct Submission
JOURNAL      Submitted (20-JUL-2004) Genoscope - Centre National de Sequencage :
             BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
             - Web : www.genoscope.cns.fr)
COMMENT      1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
             end enriched, double-strand cDNA was digested with Not I and cloned
             into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library
             was normalized. Library was constructed by Life Technologies, a
             division of Invitrogen.
FEATURES     Location/Qualifiers
     source  1..1320
             /organism="Homo sapiens"
             /mol_type="mRNA"
             /db_xref="taxon:9606"
             /clone="CS0DI021YL06"
             /tissue_type="Placenta Cot 25-normalized"
             /plasmid="pCMVSPORT_6"

ORIGIN
Query Match     61.8%;  Score 943;  DB 3;  Length 1320;
Best Local Similarity   100.0%;  Pred. No. 4.4e-204;
Matches  943;  Conservative   0;  Mismatches   0;  Indels   0;  Gaps   0;
```

```
Qy  585 GGTGGAGATCCTGGAGGGCCGACCCACCGAGTGTGTGCTGCAACCTGCGGGGGCCGGACGCG 644
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    5 GGTGGAGATCCTGGAGGGCCGACCCACCGAGTGTGTGCTGCAACCTGCGGGGGCCGGACGCG 64

Qy  645 CTACACCTTCGCCGTCCGGCCGCGGCCGTATGGCTGCGAGCCGAGCTTCGGCGGCTTCTGGAGCGC 704
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   65 CTACACCTTCGCCGTCCGGCCGCGGCCGTATGGCTGCGAGCCGAGCTTCGGCGGCTTCTGGAGCGC 124

Qy  705 CTGGTCGGAGCCTGTGTCGTGCTAGGCACCTGACGCTGGACCCCTCATCCTGACGCT 764
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  125 CTGGTCGGAGCCTGTGTCGTGCTGAGGCACCTGACGACCTGGACCCCCTCATCCTGACGCT 184

Qy  765 CTCCCTCATCCTGGTGCTGTCATCCTGGTGCTGACCGTGCTCGGCCGTGCTCTCCCACCG 824
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  185 CTCCCTCATCCTGGTGCTGTCATCCTGGTGCTGACCGTGCTCGGCCGTGCTCTCCCACCG 244

Qy  825 CCGGGCTCTGAAGCAGCAGAAGAGATCTGGCCTGGACGCCCAGCAGCCCAGAGACGAGTTGAAGG 884
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  245 CCGGGCTCTGAAGCAGCAGAAGAGATCTGGCCTGGACGCCCAGCAGCCCAGAGACGAGTTTGAAGG 304

Qy  885 CCTCTTCACCACCCACAAGGGTAACTTCCAGCTGTGGCTGTACCAGAATGATGGCTGCCT 944
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  305 CCTCTTCACCACCCACAAGGGTAACTTCCAGCTGTGGCTGTACCAGAATGATGGCTGCCT 364

Qy  945 GTGGTGAGCCCCTGCGCACCCCCTTCACGGAGGACCCCACCTGCTTCCCTGGAAGTCCTC 1004
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  365 GTGGTGAGCCCCCTGCGCACCCCCTTCACGGAGGACCCACCTGCTTCCCTGGAAGTCCTC 424

Qy 1005 AGAGCGCTGCTGGGAGGACGATGCAGGCAGTGGAGCCGGGACAGATGGATGAGGGGCCCCT 1064
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  425 AGAGCGCTGCTGGGAGGACGATGCAGGCAGTGGAGCCGGGACAGATGGATGAGGGGCCCCT 484

Qy 1065 GCTGGAGCCAGTGGCAGTGGGCACATGCCAGGATGCCCAGGATACCTATCTGGTGCTGGACAAATGGTT 1124
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  485 GCTGGAGCCAGTGGCAGTGGGCACATGCCAGGATACCTATCTGGTGCTGGACAAATGGTT 544

Qy 1125 GCTGCCCCCGGAACCGCCCCAGTGAGGACCTCCCAGGGGCCTCGTGCTGGCAGTGTGGACATAGT 1184
```

```
Db  545 GCTGCCCCCGGAACCGCCGCCAGTGAGGACCTCCCAGGGGCCTCGTGCTGGCAGTGTGGACATAGT 604

Qy 1185 GGGCCATGATGAAGGCTCAGAAGCATCCTCCTGCTCATCTGCTTGGCCTGCTTCTCGAAGCCCAG 1244
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  605 GGGCCATGATGAAGGCTCAGAAGCATCCTCCTGCTCATCTGCTTGGCCTGCTTCTCGAAGCCCAG 664

Qy 1245 CCCAGAGGGAGCCTCTGCTGCGTGCCAGCTTTGAGTACACTATCTGACCCCAGCTCCCAGCT 1304
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  665 CCCAGAGGGAGCCTCTGCTGCGTGCCAGCTTTGAGTACACTATCTGACCCCAGCTCCCAGCT 724

Qy 1305 CTTGCGTCCATGGACACTGTGCCCTGAGCTGCCCCCTACCCCACCCACTAAAGTACCT 1364
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  725 CTTGCGTCCATGGACACTGTGCCCTGAGCTGCCCCCTACCCCACCCACTAAAGTACCT 784

Qy 1365 GTACCTTGTGGTATCTGACTCTGGCATCTCAACTGACTCACAGTCACAGCTCAGGGACTCCCAGGG 1424
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  785 GTACCTTGTGGTATCTGACTCTGGCATCTCAACTGACTCACAGTCACAGCTCAGGGACTCCCAGGG 844

Qy 1425 AGCCCAAGGGGCTTATCCGATGGCCCCTACTCCAACCCTTATGAACAGCCTTATCCC 1484
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  845 AGCCCAAGGGGCTTATCCGATGGCCCCTACTCCAACCCTTATGAACAGCCTTATCCC 904

Qy 1485 AGCCGCTGAGCCTTCTGCCCCCCAGCTATGTGGCTTGCTTCTTAG 1527
        |||||||||||||||||||||||||||||||||||||||||||||
Db  905 AGCCGCTGAGCCTTCTGCCCCCCAGCTATGTGGCTTGCTTCTTAG 947
```

```
RESULT 11
BX448344
LOCUS       BX448344             984 bp    mRNA    linear    EST 05-MAY-2004
DEFINITION  BX448344 Homo sapiens FETAL LIVER Homo sapiens cDNA clone
            CS0DM009YC22 5-PRIME, mRNA sequence.
ACCESSION   BX448344
VERSION     BX448344.2   GI:47053095
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
ORGANISM    Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 984)
AUTHORS     Li,W.B., Gruber,C., Jessee,J. and Polayes,D.
TITLE       Full-length cDNA libraries and normalization
JOURNAL     Unpublished (2001)
COMMENT     On May 22, 2003 this sequence version replaced gi:31021921.
            Contact: Genoscope
            Genoscope - Centre National de Sequencage
            2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE
            Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
            1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
            end enriched, double-strand cDNA was digested with Not I and cloned
            into the Not I and EcoRV sites of the pCMVSPORT 6 vector. Library
            was not normalized. Library was constructed by Life Technologies, a
            division of Invitrogen.
            This sequence belongs to sequence cluster 9443.r
            For more information about this cluster, see
            http://www.genoscope.cns.fr/cdna?s=CS0AM009BB11QP1&c=9443.r.
FEATURES    Location/Qualifiers
     source 1..984
            /organism="Homo sapiens"
            /mol_type="mRNA"
            /db_xref="taxon:9606"
            /clone="CS0DM009YC22"
            /tissue_type="FETAL LIVER"
            /dev_stage="fetal"
            /clone_lib="Homo sapiens FETAL LIVER"
            /note="Organ: liver; Vector: pCMVSPORT 6; 1st strand cDNA
            was primed with a NotI-oligo(dT) primer. Five prime end
            enriched, double-strand cDNA was digested with Not I and
            cloned into the Not I and EcoRV sites of the pCMVSPORT 6
            vector. Library was not normalized."
ORIGIN
```

Left alignment

Query Match   56.6%;   Score 863.8; DB 5; Length 984;
Best Local Similarity 98.9%; Pred. No. 4.7e-186;
Matches 898; Conservative 2; Mismatches 2; Indels 3; Gaps 3;

```
Qy    1 ATGGACCACCTCGGGGCGTCCCTCTGGCCCCAGGTCGGCTCCCTTGTCTCCTGCTCGCT   60
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   79 ATGGACCACCTCGGGGCGTCCCTCTGGCCCCAGGTCGGCTCCCTTGTCTCCTGCTCGCT  138

Qy   61 GGGGCGGCCTGGGCGGCGCCCCCGCCTAACCTCCCGGACCCCAAGTTCGAGAGCAAAGCGGCC  120
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||| |||
Db  139 GGGGCGGCCTGGGCGGCGCCCCCGCCTAACCTCCCGGACCCCAAGTTCGAGAGCAAAGCGNCC  198

Qy  121 TTGCTGGCGGCCCGGCCCGGGGGCCCGAAGAGCTTCTGTGCTTCACCGAGCGGTTGGAGGACTTG  180
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  199 TTGCTGGCGGCCCGGCCCGGGGGCCCGAAGAGCTTCTGTGCTTCACCGAGCGGTTGGAGGACTTG  258

Qy  181 GTGTGTTTCTGGGAGGAAGCGGAGCGCGTGGGGCGCTGGGGCGGCCCGGCCCGGCAACTACAGCTTCTCC  240
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  259 GTGTGTTTCTGGGAGGAAGCGGAGCGCGTGGGGCGCTGGGGCGGCCCGGCCCGGCAACTACAGCTTCTCC  318

Qy  241 TACCAGCTCGAGGATGAGCCGCCTGCGCTGCACCAGGCTCCCACGCGCTCGT  300
        |||||||||||||||||||||||||||||||||||||||||||||||||||
Db  319 TACCAGCTCGAGGATGAGCCGCCTGCGCTGCACCAGGCTCCCACGCGCTCGT  378

Qy  301 GGTGCGGCTTCTGGTGTTCGCTGCCTACAGCCGACCACGTCGAGCTTCGTGTGCCCCCTA  360
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  379 GGTGCGGCTTCTGGTGTTCGCTGCCTACAGCCGACCACGTCGAGCTTCGTGTGCCCCCTA  438

Qy  361 GAGTTGCGCGTCACAGCAGCCTCCGGCGCTCCGGCGCTCCGGCGATATCACCGTGTCATCCACATCAAT  420
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  439 GAGTTGCGCGTCACAGCAGCCTCCGGCGCTCCGGCGCTCCGGCGATATCACCGTGTCATCCACATCAAT  498

Qy  421 GAAGTAGTGCTCCTAGACGCCCCCCGTGGGGCTGGTGGCGGCGGTGGCTGACGAGAGCGGC  480
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  499 GAAGTAGTGCTCCTAGACGCCCCCCGTGGGGCTGGTGGCGGCGGTGGCTGACGAGAGCGGC  558

Qy  481 CACGTAGTGTTGCGCTGGCTGCCCCGCGCGCCTGAGACACCCATGACGTCTCACATCCGCTAC  540
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  559 CACGTAGT-TTGCGCTGGCTGCCCCGCGCGCCTGAGACACCCATGACGTCTCACATCCGCTAC  617

Qy  541 GAGGTGGACGCTCTCCGCCCGGCCCCGGCAACCGGCCA-GGGAGCGTACAGAGAGGGTGGAGATCCTGGA  599
        |||||||||||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||||
Db  618 GAGGTGGACGCTCTCCGCCCGGCCCCGGCAACCGGCCACGGGAGCGTACAGAGAGGGTGGAGATCCTGGA  677

Qy  600 GGGCCGCACCAGTTGTGTGCTGAGCAACCTGGCGGGGCCGAGACGGCGCTCACCTTCGCCGT  659
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  678 GGGCCGCACCAGTTGTGTGCTGAGCAACCTGGCGGGGCCGAGACGGCGCTCACCTTCGCCGT  737

Qy  660 CCGCGCGGCCTATGCTGAGCGCAGCTTCGGCGGCCTTCGGAGCGCCTGT  719
        |||||||||||||||||||||||||||||||||||||||||||||||
Db  738 CCCGCGCGGCCTATGCTGAGCGCAGCTTCGGCGGCCTTCGGAGCGCCTGT  797

Qy  720 GTCCTGCTGACGCTAGGGACCTGGACCCCCTCATCCTGACGCGTCTCCCTCATCCTCGT  779
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  798 GTCGCTGCTGACGCTAGGGACCTGGACCCCCTCATCCTGACGCGTCTCCCTCATCCTCGT  857

Qy  780 GGTCATCCTGGTGCTGCTGACCGTGTCCGGCGCTGTCTCCCACCGCGGCCTCTGAAGCA  839
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  858 GGTCATCCTGGTGCTGCTGACCGTGTCCGGCGCTGTCTCCCACCGCGGCCTCTGAAGCA  917

Qy  840 GAAGATCTGCCTGCCTGCATCCCGAGCCCAGAGAGCCGAGTTGAAGGCCTCTTCACCACCCA  899
        ||:||||||||:|| |||:||||||||||||||||||||||||||||| :||:||||||||
Db  918 GAAGATCTGCGCCTTGGATCGCGAGCGCCAGAGAGCGCGAGTTTGAAGG-CTYTTCACCACCAM  976

Qy  900 CAAGGGTA  907
        ||||||||
Db  977 CAAGGGTA  984
```

Right column (ORIGIN / Accession)

Query Match   56.1%;   Score 857.4; DB 1; Length 1079;
Best Local Similarity 96.4%; Pred. No. 1.4e-184;
Matches 909; Conservative 9; Mismatches 20; Indels 5; Gaps 4;

```
Qy    1 ATGGACCACCTCGGGGCGTCCCTCTGGCCCCAGGTCGGCTCCCTTGTCTCCTGCTCGCT   60
        ||||||||||||||||||||||||||||||||||||||||||||||||| |||||||||
Db   43 ATGGACCACCTCGGGGCGTCCCTCTGGCCCCAGGTCGGCTCCCTTGTCCCCTGCTCGCT  102

Qy   61 GGGGCGGCCTGGGCGGCGCCCCCGCCTAACCTCCCGGACCCCAAGTTCGAGAGCAAAGCGGCC  120
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  103 GGGGCGGCCTGGGCGGCGCCCCCGCCTAACCTCCCGGACCCCAAGTTCGAGAGCAAAGCGGCC  162

Qy  121 TTGCTGGCGGCCCGGCCCGGGGGCCCGAAGAGCTTCTGTGCTTCACCGAGCGGTTGGAGGACTTG  180
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  163 CTGCCGGCGGCCCGGCCCGGGGGCCCGAAGAGCTTCTGTGCTTCACCGAGCGGTTGGAGGACTTG  222

Qy  181 GTGTGTTTCTGGGAGGAAGCGGAGCGGCGGCCGCCGGGTGGGCCGGTGGCCCGGCAACTACAGCTTCTCC  240
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  223 GTGTGTTTCTGGGAGGAAGCGGAGCGGCGGCCGCCGGGTGGGCCGGTGGCCCGGCAACTACAGCTTCTCC  282

Qy  241 TACCAGCTCGAGGATGAGCCGCCTGCGCTGCACCAGGCTCCGACACGTCGT  300
        |||||||||||||||||||||||||||||||||||||||||||||||||||
Db  283 TACCAGCTCGAGGATGAGCCGCCTGCGCTGCACCAGGCTCCGACACGTCGT  342

Qy  301 GGTGCGGCGGTTGCGCTTCTCGTGTTCGCTGCCTACAGCCGCGACGTCGTCGAGCTTCGTGTGCCCCCTA  360
        |||||| ||||||||||||||||||||||||||||||||||||||||| ||||||||||||| |||||
Db  343 GGTGCGGCGGTTGCGCTTCTCGTGTTCGGT-GTTCGCTGCGTTCACACGCGACACGTCGAGCTTCGTGTGCCCCCWA  402

Qy  361 GAGTTGCGCGTCACAGCAGCCTCCGGCGCTCCGGCGCTCCGGCGATATCACCGTGTCATCCACATCAAT  420
        |||||||||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||
Db  403 GAGTTGCGCGTCACAGCAGCCTCCGGCGCTCCGGCGCTCCGGCGAWATCACCGTGTCATCCACATCAAT  462

Qy  421 GAAGTAGTGCTCCTAGACCCCCCGTGGGGCTGGTGGCGGCGGTGACGAGAGACGCGG  479
        ||||||||||||||||||| |||||||||||||||||||| |||||||||||||
Db  421 ...
```

RESULT 12
AL553257
LOCUS   AL553257   1079 bp   mRNA   linear   EST 30-MAR-2004
DEFINITION AL553257 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens cDNA

```
Db   463 GAAGTAGTGCTCCMAGACGCCCCCGTGGGCGGCGGTTGGCGCGGTTGGCTGACGAGAGCGG 522
             |||||||||||||||||||||||||||||||||||||||||||||||||
Qy   480 CCACGTAGTGTTGCGCTGGCTGCCCGCCGCCTGAGACACCCATGACGTCTCACATCCGCTA 539
             |||||||||||||||||||||||||||||||||||||||||||||||||
Db   523 CCACGTAGTTGCGCTGGCTGCCCGCCGCCTGACACACCCATGACGTCTCACATCCGCCA 582

Qy   540 CGAGGTGGACGTCTCGGCCGCCGGCAACCGGCGCGCAGGGAGCGTACAGAGGGTGGAGATCCTGGA 599
             |||||||||||||||||||||||||||||||||||||||||||||||||
Db   583 CGAGGTGGACGTCTCGGCCGCCGGCAACCGGCGCGCAGGGAGCGTACAGAGGGTGGAGATCCTGGA 642

Qy   600 GGGCCGCACCGAGTGTGTGCTGAGCAACCTGCGGGGCCGGACCGCGCGCTACACCTTCGCCGT 659
             |||||||||||||||||||||||||||||||||||||||||||||||||
Db   643 GGGCCGCACCGAGTGTGTGCTGAGCAACCTGCGGGGCCGGACCGCGCGCTACACCTTCGCCGT 702

Qy   660 CCGCGCGGCCAGTGTGCTGAGCCAGCTTCGGCGGCGCCTGGTCGCGGAGCCTGT 719
             |||||||||||||||||||||||||||||||||||||||||||||||||
Db   703 CCGCGCGGCCAGTGTGCTGAGCCAGCTTCGGCGGCGCCTGGTCGCGGAGCCTGT 762

Qy   720 GTCGCTGCTGACGACCTGGCCTAGCGACCTGGACCCCCTATCCTGACGCTCTCCCATCCTCGT 779
             |||||||||||||||||||||||||||||||||||||||||||||||||
Db   763 GTCGCTGCTGACGCGCCTAGCGACCTGGACCCCCTATCCTCGACGCTCTCCCATCCTCGT 822

Qy   780 GGTCATCCTGGTGCTGCTGACACCGTGCTGCTCTCCCCACCCGGGCTCTGAAGCA 839
             |||||||||||||||||||||||||||||||||||||||||||||||||
Db   823 GGTCATCCTGGTGCTGCTGACACCGTGCTGCTCTCCCCACCCGGGCTCTGAAGCA 882

Qy   840 GAAGATCTGGCCTGGCCTGGCCTAGCCAGAGAGGAGTTTGAAGGCCTCTTCACCACCCA 899
             |||||||||||||||    ||||||||||:|||||| ||||| | :: |||
Db   883 GAAGATCTGGCCTGGCATCYCGAGCCCAGAGCCGAGTTKAAAGGC--TCTTCACMACCA 940

Qy   900 CAAGGGTAACTTCCAGCTGTGGCTGTTACCAGAATGATGGCTGC 942
             |||||||||||||||||||:  ||||||||||  ||||||||::||
Db   941 CAAGGGTAAC-TCCAGCTGTKGCTGTA-CARAATGATGGCKSC 981
```

```
RESULT 13
BX335579
LOCUS        BX335579                1005 bp    mRNA        linear   EST 08-APR-2004
DEFINITION   BX335579 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens cDNA
             clone CS0DI016YL05 5-PRIME, mRNA sequence.
ACCESSION    BX335579
VERSION      BX335579.2  GI:46281994
KEYWORDS     EST.
SOURCE       Homo sapiens (human)
ORGANISM     Homo sapiens
             Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
             Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE    1  (bases 1 to 1005)
AUTHORS      Li,W.B., Gruber,C., Jessee,J. and Polayes,D.
TITLE        Full-length cDNA libraries and normalization
JOURNAL      Unpublished (2001)
COMMENT      On May 2, 2003 this sequence version replaced gi:30345434.
             Contact: Genoscope
             Genoscope - Centre National de Sequencage
             2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE
             Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
             1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
             end enriched, double-strand cDNA was digested with Not I and cloned
             into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library
             was normalized. Library was constructed by Life Technologies, a
             division of Invitrogen. This sequence belongs to sequence cluster
             9443.r
             For more information about this cluster, see
             http://www.genoscope.cns.fr/cdna?s=CS0DI016CF03QPlkc=9443.r.
FEATURES             Location/Qualifiers
     source          1..1005
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="CS0DI016YL05"
                     /tissue_type="PLACENTA COT 25-NORMALIZED"
                     /clone_lib="Homo sapiens PLACENTA COT 25-NORMALIZED"
                     /note="1st strand cDNA was primed with a NotI-oligo(dT)
                     primer. Five prime end enriched, double-strand cDNA was
                     digested with Not I and cloned into the Not I and EcoR V
                     sites of the pCMVSPORT 6 vector. Library was normalized."
```

```
ORIGIN

Query Match        53.2%;   Score 812.2;  DB 5;   Length 1005;
Best Local Similarity 96.1%;   Pred. No. 2.7e-174;
Matches  864;  Conservative  18;  Mismatches  12;  Indels  5;  Gaps  5;

Qy     1 ATGGACCACCTCGGGCGCTCCCTCCTGCCCCAGGTCGGCTCCCTTGTCTCCTGCTCGCT 60
             |||||||||||||||||||||||||||||||||||||||||||||||||
Db    93 ATGGACCACCTCGGGCGCTCCCTCCTGCCCCAGGTCGGCTCCCTTGTCTCCTGCTCGCT 152

Qy    61 GGGGCGCCGCCTGGCGCGCCCCCGCCTAACCTCCCGGACCCCAAGTTCGAGAGCAAAGCGGCC 120
             |||||||||||||||||||||||||||||||||||||||||||||||||
Db   153 GGGGCGCCGCCTGGCGCGCCCCCGCCTAACCTCCCGGACCCCAAGTTCGAGAGCAAAGCGGCC 212

Qy   121 TTGCTGCGGCCCGGGGCCCGAAGAGCTTCTGTGTTCACCGAGCGGTTGGAGGACTTG 180
             |||||||||||||||||||||||||||||||||||||||||||||||||
Db   213 TTGCTGCGGCCCGGGGCCCGAAGAGCTTCTGTGTTCACCGAGCGGTTGGAGGACTTG 272

Qy   181 GTGTGTTTCTGGAGGAAGCGGGAGCGCGTGGGGTGGCCCGGCAACTACAGCTTCTCC 240
             |||||||||||||||||||||||||||||||||||||||||||||||||
Db   273 GTGTGTTTCTGGAGGAAGCGGGAGCGCGTGGGGTGGCCCGGCAACTACAGCTTCTCC 332

Qy   241 TACCAGCTGGAGGATGAGCCAGCTCGCGCTCGCCTGTGTCGCGGAGCTCCCACGGCTCGT 300
             |||||||||||||||||||||||||||||||||||||||||||||||||
Db   333 TACCAGCTGGAGGATGAGCCAGCTCGCGCTCGCCTGTGTCGCGGAGCTCCCACGGCTCGT 392

Qy   301 GGTGGCGGTGGCTTCTGGTGTTCGCTCACAGCCGACACGTCGAGCTTCGTGCCCCTA 360
             |||||||||||||||||||||||||||||||||||||||||||||||||
Db   393 GGTGGCGGTGGCTTCTGGTGTTCGCTCACAGCCGACACGTCGAGCTTCGTGCCCCTA 452

Qy   361 GAGTTGCGCGGTCACAGCAGCGTCCCGCGCTCCGCGATATCACCGTGTCATCCACATCAAT 420
             |||||||||||||||||||||||||||||||||||||||||||||||||
Db   453 GAGTTGCGCGGTCACAGCAGCGTCCCGCGCTCCGCGATATCACCGTGTCATCCACATCAAT 512

Qy   421 GAAGTAGTGCTCCTAGACGCCCGTGGGCTGGTGGCTGGCGGGC 480
             |||||||||||||||||||||||||||||||||||||||||||||
Db   513 GAAGTAGTGCTCCTAGACGCCCGTGGGGCTGGTGGCTGGCGGGC 572

Qy   481 CACGTAGTGTTGCGCTGGCTGCCCTGAGACACCCATGACGTCTCACATCCGCTAC 540
             |||||||||||||||||||||||||||||||||||||||||||||||||
Db   573 CACGTAGTTATGCGCTGGCTGGCTGCCCTGAGACACCCATGACGTCTCACATCCGCTAC 632

Qy   541 GAGGTGGACGTCTCGGCCGCCAACCGGCAGGGAGCGCAGGGAGCGTACAGAGGGTGGAGATCCTGGAG 600
             |||||||||||||||||||||||||||||||||||||||||||||||||
Db   633 GAGGTGGACGTCTCGGCCGCCAAC-GCGCAGGGAGCGTACAGAGGGTGGAGATCCTGGAG 691

Qy   601 GGCCGCACCGAGTGTGTGCTGAGCAACCTGCGGGGCCGGACCGCGCGCTACACCTTCGCCGTC 660
             |||||||||||||||||||||||||||||||||||||||||||||||||
Db   692 GGCCGCACCGAGTGTGTGCTGAGCAACCTGCGGGGCCGGACCGCGCGCTACACCTTCGCCGTC 751

Qy   661 CGCGCGGCCAGTGTGCTGAGCCAGCTTCGGCGGCGCCTTCGGAGCCGGTCGTGGTCGCGGAGCCGTG 720
             |||||||||||||||||||||||||||||||||||||||||||||||||
Db   752 CGCGCGGCCAGTGTGCTGAGCCAGCTTCGGCGGCGCCTTCGGAGCCGTCGGTCGCGGAGCCGTG 811

Qy   721 TCGCTGCTGACGCGCCTAGCGACCTGGA---CCCCTATCCTGACGCTCTCCCATCCTCGTG 780
             |||||||||||||||||||||||||| |:||||||||||||||||||
Db   812 TCGCTGCTGACGCGCCTAGCGACCTGGA-CCCCTATCCTCGACGCTCTCCCATCCTCGTG 870

Qy   781 GTCATCCTGGTGCTGCTGACGCGCATCCGAGCCCGTGCTCGCCTGCTCTCCCACCCGGGCTCTGAAGCAG 840
             |:||||||||||||||||||||||||| |||||||||||||||||||||:
Db   871 GKCATCCTGGTGCTGCTGACGCGCTCTSG-GCTGCTCTCCCACCCGCGGGSTCTGAGVAG 929

Qy   841 AAGATCTGGCCTGGACATCCGAGCCCAGAGCAGAGGCGAGTTTGAAGGCCTCTTCACCACCCA 899
             || |:||:|:|:|||| :|||| : ||: |:|| ||
Db   930 -AGRTTTKGSCTKGSATYCCGAG--CCAGARARCGAGTTTAARGCYYYTYACCACMCAA 986
```

```
RESULT 14
BX448343/c
LOCUS        BX448343                1051 bp    mRNA        linear   EST 05-MAY-2004
```

```
DEFINITION  BX448343 Homo sapiens FETAL LIVER Homo sapiens cDNA clone
            CS0DM009YC22 3-PRIME, mRNA sequence.
ACCESSION   BX448343
VERSION     BX448343.2  GI:47052187
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
ORGANISM    Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 1051)
AUTHORS     Li,W.B., Gruber,C., Jessee,J. and Polayes,D.
TITLE       Full-length cDNA libraries and normalization
JOURNAL     Unpublished (2001)
COMMENT     On May 22, 2003 this sequence version replaced gi:31019923.
            Contact: Genoscope
            Genoscope - Centre National de Sequencage
            2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE
            Email: segref@genoscope.cns.fr, Web : www.genoscope.cns.fr
            1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
            end enriched, double-strand cDNA was digested with Not I and cloned
            into the Not I and EcoRV sites of the pCMVSPORT 6 vector. Library
            was not normalized. Library was constructed by Life Technologies, a
            division of Invitrogen.
            This sequence belongs to sequence cluster 9443.r
            For more information about this cluster, see
            http://www.genoscope.cns.fr/cdna?rs=CS0AM009BB1NP1&c=9443.r.
FEATURES    Location/Qualifiers
    source  1..1051
            /organism="Homo sapiens"
            /mol_type="mRNA"
            /db_xref="taxon:9606"
            /clone="CS0DM009YC22"
            /tissue_type="FETAL LIVER"
            /dev_stage="fetal"
            /clone_lib="Homo sapiens FETAL LIVER"
            /note="Organ: liver; Vector: pCMVSPORT 6; 1st strand cDNA
            was primed with a NotI-oligo(dT) primer. Five prime end
            enriched, double-strand cDNA was digested with Not I and
            cloned into the Not I and EcoRV sites of the pCMVSPORT 6
            vector. Library was not normalized."
ORIGIN

Query Match        51.9%;  Score 792.2;  DB 5;  Length 1051;
Best Local Similarity  97.7%;  Pred. No. 9.9e-170;
Matches  820;  Conservative  4;  Mismatches  12;  Indels  3;  Gaps  2;

Qy   689 GCGGGCTTCTGGAGCGCCTGGTCGGAGCCTGGTCGCTGCTGACGCGCTGGTAGCGGACGGAGCCTGGACC 748
         |||||||| |||| ||| |||||| ||||||||||||||||| :||||||||||         |||||||
Db  1042 GCGGGCTTTKGAGGCGGCCTGGTGT--GCTGGTGTCGCTGGT--ACSCTAGGGACCTGGACC 986

Qy   749 CCCTCATCCTGACGCGTCTCCCCTCATCCTCGTGGTCATCCTGTGCGTGCTGTCGACCGTGGTGCTCTCG 808
         |||||||||||| ||| ||| |||||||||| |||||||||| |||||||||||||||||||||||
Db   985 CCCTCATCCTAACGTTCTCCCCTCATCCTCGTGGTGTCATCCTGTGCGTGCTGGCTGCTCGACCTGTGCTG 926

Qy   809 CGCTGCTCCCACCGCCGGCCGCGGCGCTCTGAAGCAGCAGAAGATCTGGCTCGCTGCATCCGAGCCCCAG 868
         ||||||||| |:||||| |||||||||||||||||||||||||||||||||||:||||||||||||||||
Db   925 CGCTGCTCCCACCGCCGGCCGCGGCGCTCTGAAGCAGCAGAAGATCTGGCTCGCTGCATCCGAGCCCCAG 866

Qy   869 AGAGGAGTTTGAAGGCCTCTTCACCACCACAAGGTAACTTCCAGCTGTGGCTGTACC 928
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   865 AGAGGAGTTTGAAGGCCTCTTCACCACCACAAGGTAACTTCCAGCTGTGGCTGTACC 806

Qy   929 AGAATGATGGCGCTGCCTGTGGTGGAGCGCGTGGGACGCCCCTGCCCCCTCACGGAGGACACCACCTGCTT 988
         |||||||||||| ||||||| ||||||||||  ||||||||||||||||||||||||||||||||||||||
Db   805 AGAATGATGGCGCTGCCTGTGGTGGAGCGCGCCCTGCACCCCCTTCACGGAGGACCACCACCACCTGCTT 746

Qy   989 CCCTGGAAGTCCTCTCAGAGGCGCTGCTGGGGGACCAGCGCTGCTGGTGAGGCAGTGGAGCCGGGGACAG 1048
         |||||||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||||||||
Db   745 CCCTGGAAGTCCTCTCAGAGGCGCTGCTGGGGGACCAGCGCTGCTGGTGAGGCAGTGGAGCCGGGGACAG 686

Qy  1049 ATGATGAGGGCCCCCTGCTGGAGCCAGTGGGCAGTGGGCAGTGAGCATGCCCAGGATACCTATCTGG 1108
         |||||||||| ||||||||| ||||||||||||||||||||||||| ||| |||||| ||||||||||
Db   685 ATGATGAGGGCGCCCCCTGCTGGAGCCAGTGGGCAGTGGGCAGTGAGCATGCCCAGGATACCTATCTGG 626
```

```
RESULT 15
AY414847
LOCUS       AY414847               1115 bp    DNA     linear   GSS 17-DEC-2003
DEFINITION  Pan troglodytes EPOR gene, VIRTUAL TRANSCRIPT, partial sequence,
            genomic survey sequence.
ACCESSION   AY414847
VERSION     AY414847.1  GI:39770806
KEYWORDS    GSS.
SOURCE      Pan troglodytes (chimpanzee)
ORGANISM    Pan troglodytes
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pan.
REFERENCE   1 (bases 1 to 1115)
AUTHORS     Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A.,
            Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
            Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
            Adams,M.D. and Cargill,M.
TITLE       Inferring nonneutral evolution from human-chimp-mouse orthologous
            gene trios
JOURNAL     Science 302 (5652), 1960-1963 (2003)
PUBMED      14671302
REFERENCE   2 (bases 1 to 1115)
AUTHORS     Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A.,
            Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
            Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
            Adams,M.D. and Cargill,M.
TITLE       Direct Submission
JOURNAL     Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,
            Rockville, MD 20850, USA
COMMENT     This sequence was made by sequencing genomic exons and ordering
            them based on alignment.
FEATURES    Location/Qualifiers
    source  1..1115
            /organism="Pan troglodytes"
            /mol_type="genomic DNA"
            /db_xref="taxon:9598"
    gene    <1..>1115
            /gene="EPOR"
            /locus_tag="HCM360"
ORIGIN

Query Match        51.7%;  Score 789;  DB 9;  Length 1115;
Best Local Similarity  71.0%;  Pred. No. 5.4e-169;

Qy  1109 TGCTGGACAAATGGTTGCTGCCCCGGAACCCGGCCTCCAGTGAGGCCTGGTG 1168
         |||||||||||||||||||||||||||||||||||||||||||||||||||
Db   625 TGCTGGACAAATGGTTGCTGCCCCGGAACCCGGCCTCCAGTGAGGCCTGGTG 566

Qy  1169 GCAGTGTGGACATAGTGGCCATGAGGCTCAGAAGGCTCAGAAGCATCTCCTGCTGCTCATCTGCTT 1228
         |||||||||||||||||||||||||||         |||||||||||||||||||||||||||||
Db   565 GCAGTGTGGACATAGTGGCCATGAGGATGAAGGCTCAGAAGCATCTCCTGCTGCTCATCTGCTT 506

Qy  1229 TGGCCTCGAAGCCCAGACGAGGAGCCTCTGCTGCCAGCTTTGAGTACACTATCCTGG 1288
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   505 TGGCCTCGAAGCCCAGACGAGGAGCCTCTGCTGCCAGCTTTGAGTACACTATCCTGG 446

Qy  1289 ACCCAGCTCCCAGCTCTTGCGTCCATGGACACTGTGCCCTGAGCTGCCCCGAGCTGCCCCTACCCCAC 1348
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   445 ACCCAGCTCCCAGCTCTTGCGTCCATGGACACTGTGCCCTGAGCTGCCCCGAGCTGCCCCTACCCCAC 386

Qy  1349 CCCACCTAAAGTACCTGTACCTTGTGTATCTGACTCTGGCATCTCAACTGACTACAGCT 1408
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   385 CCCACCTAAAGTACCTGTACCTTGTGTATCTGACTCTGGCATCTCAACTGACTACAGCT 326

Qy  1409 CAGGGGACTCCCAGGGAGCCCAAGGGGCTTATCCGATGGCCCTACTCCAACCCTTATG 1468
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   325 CAGGGGACTCCCAGGGAGCCCAAGGGGCTTATCCGATGGCCCTACTCCAACCCTTATG 266

Qy  1469 AGAACAGCCTTATCCCAGCCGCTGAGCCTCTGCCCCCCAGCTATGTGGCTTGCTTCTTAG 1527
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   265 AGAACAGCCTTATCCCAGCCGCTGAGCCTCTGCCCCCCAGCTATGTGGCTTGCTTCTTAG 207
```

Matches 792;   Conservative   0;   Mismatches 323;   Indels   0;   Gaps   0;

```
Qy    116 CGGCCTTGCTGGCGGCCCCGGGGCCCGAGAGCTTCTGTGCTTCACCGAGCGGTTGGAGG  175
          ||||||||||||||||||||||||||||||||||| |||||||||| |||||||||||
Db      1 CGGCCTTGCTGGCGGCCCCGGGGCCCGAACAGCTTCTGTGCTTCGTGCTTCACCGAGAGGTTGGAGG   60

Qy    176 ACTTGGTGTGTTTCTGGGAGGAAGCGGCGAGCCGCTGGGGTGGGCCAACTACAGCT  235
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61 ACTTGGTGTGTTTCTGGGAGGAAGCGGCGAGCCGCTGGGGTGGGCCAACTACAGCT  120

Qy    236 TCTCCTACCAGCTCGAGGATGAGCCATGAGCAAGTGTGTCGCCTCGCCAGGCTCCCACGG  295
          |||||||||||||||||||
Db    121 TCTCCTACCAGCTCGANNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN  180

Qy    296 CTCGTGGTGCGGTCGCGTTCTGGTGTTCGCTGCCTACAGCCGACACGTCGAGCTTCGTGC  355
          NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN
Db    181 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN  240

Qy    356 CCCTAGAGTTGCGCGTCACAGCAGCCTCCGGCGCTCCGCGATATCACCGTGTCATCCACA  415
          |||||||||||||||        |||||||||||||| ||||||||||| |||
Db    241 NNNNNNNNNNNNNNNGCAGCCTCCGGCGCTCCGCGATATCACCNNGTCATCCACA  300

Qy    416 TCAATGAACTAGTCCTCTAGACGCCCCCCGTGGGGGCTGGTTGGCTGACGAGA  475
          ||||||||||||||||| ||||||||||||||||| ||||||||||
Db    301 TCNNNNNNNNNNNCTCCTAGACGCCCCCCGTGGGNCTGGTTGGCTGACGAGA  360

Qy    476 GCGGCCACGTAGTGTTGCGGCTGGCTCCCCGCCTGAGACACCCATGACGTCTCACATCC  535
          ||||||||||||||||||         ||||||||||||||
Db    361 GCGGCCACGTAGTGTTGCNNNNGGCTCCCGCCGNNNGAGACACCCATGACGTCCCACATCC  420

Qy    536 GCTACGAGGTGACGCTCTCGGCCGCCGGCAACGGCCCAGGGAGCCGTACAGAGGGTGGAGATCC  595
          |||| |||||||||||||        |||||||||||||||||| ||||||||
Db    421 GCTACNAGGTGACGNNNNCGGCCGGCAGGGANCGTACAGAGGNNNNNNNN  480

Qy    596 TGGAGGGCCGCACCGAGTGTGTCTGAGCAACCTGCGGGGCCCGACGCGCTACACCTTCG  655
          NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN
Db    481 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN  540

Qy    656 CCGTCCGCGCGGCCGTATGCGCTGAGCCGACCTTCGGCGGCTTCTCGGAGCGCCTGTCGCGAGC  715
          NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN
Db    541 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN  600

Qy    716 CTGTGTCGCTGCTGACGCCTAGCGACCTGGACCCCCTCATCGTGACGCTCTCCCTCATCC  775
          |||||||||||||||||||||||||||
Db    601 NNNNNNNNNNNNNNNNNGACCCCCTCATCCTGACGCTCTCCCTCATCC  660

Qy    776 TCGTGGTCATCCTGGTCGTGCTGCTCGACCGGCTGCTCTCCCACCGGCTCTGA  835
          |||| |||||||||||||||||||||||||||||||||||||||  ||
Db    661 TCGTGCTTCATCCTGGTCGTGCTGCTCGACCGGCTGCTCTCCCACCGGCGNNNNNNNGA  720

Qy    836 AGCAGAAGATCTGGCCTGCATCCTGCCCAGAGCCCAGAGCCGAGTTTGAAGGCCTCTTCACCA  895
          |||||||    ||||||||||||||||||||||||||||||||||||||
Db    721 AGCAGAAGANNNNGGCCTGCATCCCGAGCCCAGAGCCGAGTTTGAAGGCCTCTTCACCA  780

Qy    896 CCCACAAGGGTAACTTCACGGAGGAGACCACCTGGTCGTGTACCAGAATGATGGCCTGTGTGGTGGAGCC  955
          |||||||||||||||||||||||||||||||||||||||||||||||||||
Db    781 CCCACAAGGGTAACTTCACGGAGGAGACCACCCTGGTGTACCAGAATGATGGCCTGTGTGTGGAGCC  840

Qy    956 CCTGCACCCCCTTCACGGAGGACCACCTGCTTCCTGGAAGTCCTCTCAGAGCGCTGCT  1015
          |||||||||||||||||||||||||||||||||||||||||
Db    841 CCTGCACCCCCTTCACGGAGGACCACCTGCTTCCTGGAAGTCCTCTCAGAGCGCTGCT  900

Qy   1016 GGGGGACGATGCAGTGGCAGTGGAGCGCGGGGACAGATGATGAGGCGCCCCCTGCTGGAGCCAG  1075
          |||||||||||||||||||||||||||||||||||||||||||||||||||
Db    901 GGGGGACNATGCAGTGGCAGTGGAGCGCGGGGACAGATGATGAGGGCCCCCCTGCTGGAGCCAG  960

Qy   1076 TGGGCAGTGAGCATGCCCAGGATACCTATCTGCTGGACAAATGGTTGCTGCCCCCGGA  1135
          |||||||||||||||||||||||||| ||| |||||||||||||
Db    961 TGGGCAGTGAGCATGCCCAGGATACCTATCTGCTGGACAAATGGTTGCTGCCCCCAGA  1020

Qy   1136 ACCCGCCCAGTGAGGACCTCCCAGGGCCTGGTGCAGTGGCAGTAGTGGCCATGGATG  1195
          |||||||||||||||||||||||||||||||||||||||||||||||||
Db   1021 ACCCGCCCAGTGAGGACCTCCCAGGGCCTGGTGCAGTGGCAGTAGTGGCCATGGATG  1080
```

```
Qy   1196 AAGGCTCAGAAGCATCCTCCTGCTCATCTGCTTTG  1230
          |||||||||||||||||||||||||||||||||||
Db   1081 AAGGCTCAGAAGCATCCTCCTGCTCATCTGCTTTG  1115
```

Search completed: April 21, 2005, 04:08:44
Job time : 5271 secs

```
                GenCore version 5.1.6
        Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on:      April 20, 2005, 19:27:26 ; Search time 849.5 Seconds
                                          (without alignments)
                                       10640.901 Million cell updates/sec

Title:          US-09-016-159D-4
Perfect score:  1527
Sequence:       1 atggaccacctcgggggcgtc.........gctatgtggctgctgctcttag 1527

Scoring table:  IDENTITY NUC
                Gapop 10.0 , Gapext 1.0

Searched:       4390206 seqs, 2959870667 residues

Total number of hits satisfying chosen parameters:    8780412

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :  N_Geneseq_16Dec04.*
         1: geneseqn1980s:*
         2: geneseqn1990s:*
         3: geneseqn2000s:*
         4: geneseqn2001as:*
         5: geneseqn2001bs:*
         6: geneseqn2002as:*
         7: geneseqn2002bs:*
         8: geneseqn2003as:*
         9: geneseqn2003bs:*
        10: geneseqn2003cs:*
        11: geneseqn2003ds:*
        12: geneseqn2004as:*
        13: geneseqn2004bs:*
```

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	1527	100.0	1527	2	AAQ82990	Aaq82990 Human ery
2	1527	100.0	1848	13	ADQ83546	Adq83546 Human tum
3	1527	100.0	1849	12	ADO05723	Ado05723 Human ery
4	1527	100.0	1849	12	ADP10351	Adp10351 Reference
5	1527	100.0	1849	13	ACN40465	Acn40465 Tumour-as
6	1527	100.0	1883	2	AAQ05748	Aaq05748 EPO recep
7	1525.4	99.9	1527	6	ABL51519	Abl51519 Human ery
8	1525.4	99.9	1818	2	AAQ81892	Aaq81892 Human ery
9	1520.6	99.6	1585	10	ADE28676	Ade28676 Human NOV
10	1520.6	99.6	1585	12	ADM93421	Adm93421 Human NOV
11	1515.8	99.3	1818	2	AAQ53995	Aaq53995 Human EPO
12	1312.2	85.9	1317	3	AAZ49634	Aaz49634 Truncated
13	1307.4	85.6	1317	2	AAZ49636	Aaz49636 Mutant R1
14	1209	79.2	1733	10	ADE28670	Ade28670 Human NOV
15	1209	79.2	1733	10	ADE28672	Ade28672 Human NOV
16	1209	79.2	1733	12	ADM93417	Adm93417 Human NOV
17	1209	79.2	1733	12	ADM93415	Adm93415 Human NOV
18	1204.2	78.9	1435	12	ADE28674	Ade28674 Human NOV
19	1204.2	78.9	1435	12	ADM93419	Adm93419 Human NOV
20	1080.6	70.8	1740	2	AAQ05747	Aaq05747 EPO recep
21	1075.8	70.5	1741	2	AAQ53994	Aaq53994 MEL EPO r
22	1075.8	70.5	1741	2	AAQ44854	Aaq44854 Mouse sol
23	1075.8	70.5	1741	2	AAQ81891	Aaq81891 Mouse ery
24	1024	67.1	1024	10	ABZ83469	Abz83469 Toxicolog
25	811.2	53.1	5068	13	ADR70427	Adr70427 Vector pA
26	811.2	53.1	5068	13	ADR70421	Adr70421 Vector pA
27	810.4	53.1	5683	6	AAL41116	Aal41116 plasmid p
28	810.4	53.1	6122	6	ABN84078	Abn84078 Plasmid p
29	810.4	53.1	6122	6	AAL41112	Aal41112 Plasmid p
30	810.4	53.1	6122	6	AAL41115	Aal41115 Plasmid p
31	810	53.0	1121	8	ABZ23425	Abz23425 Nucleotid
32	810	53.0	4896	13	ADR70403	Adr70403 Vector pA
33	809.6	53.0	1056	8	ABZ23429	Abz23429 Nucleotid
34	809.6	53.0	5059	13	ADR70409	Adr70409 Vector pA
35	809.2	53.0	1131	8	ABZ23424	Abz23424 Nucleotid
36	808.8	53.0	6125	6	AAL41114	Aal41114 Plasmid p
37	808	52.9	1059	8	ABZ23428	Abz23428 Nucleotid
38	808	52.9	5059	13	ADR70415	Adr70415 Vector pA
39	750.4	49.1	4990	2	AAT48800	Aat48800 Plasmid m
40	750	49.1	750	2	AAV04434	Aav04434 Erythropo
41	750	49.1	750	2	AAZ30854	Aaz30854 Human ery
42	748.2	49.0	6256	6	AAL43173	Aal43173 pCAGGS-hS
43	741	48.5	5565	6	AAL43171	Aal43171 Human exp
44	700.4	45.9	1666	2	AAX58148	Aax58148 CadC-fusi
45	602.8	39.5	4883	6	ABQ73377	Abq73377 Vector pR

ALIGNMENTS

```
RESULT 1
AAQ82990
ID   AAQ82990 standard; cDNA; 1527 BP.
XX
AC   AAQ82990;
XX
DT   25-MAR-2003  (revised)
DT   07-OCT-1995  (first entry)
XX
DE   Human erythropoietin receptor cDNA.
XX
KW   Erythropoietin receptor; extracellular domain; plasmid pJYL26; ds.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   CDS             1..1527
FT                   /*tag= a
FT   misc_feature    25..250
FT                   /*tag= c
FT                   /note= "extracellular domain"
FT   misc_feature    73..750
FT                   /*tag= d
FT                   /note= "extracellular domain"
FT   misc_feature    73..87
FT                   /*tag= b
FT                   /note= "forward primer AAQ82991 site"
XX
PN   WO9505469-A1.
XX
PD   23-FEB-1995.
XX
PF   15-AUG-1994;   94WO-US009298.
XX
PR   16-AUG-1993;   93US-00106815.
XX
PA   (LEBJ/) LEE J Y.
XX
PI   Lee JY;
XX
DR   WPI; 1995-098767/13.
DR   P-PSDB; AAR70032.
XX
```

PT New pure human erythropoietin receptor fragment - obtd. by expression as
PT a fusion protein having a thrombin proteolytic cleavage site.
XX
PS Disclosure; Page 27-29; 42pp; English.
XX
CC The cDNA encodes human full-length erythropoietin receptor (EPO-R).
CC Primers AAQ82991 and AAQ82292 are used to isolate fragments (nt 73-750
CC and 25-250) representing the extracellular domain of the receptor. The
CC fragments are cloned in vector plasmid pGEX-2T, resulting in vector
CC pJYL26, which encodes the extracellular domain as a fusion protein with
CC glutathione-S-transferase. The extracellular domain of EPO-R is used for
CC investigating the structure of the EPO-R and for identifying factors
CC involved in regulating differentiation and proliferation mechanisms in
CC erythroid progenitor cells. It can also be used for identifying and
CC quantitating EPO and EPO-R as well as in understanding haematopoietic
CC malignancy and some cardiovascular system disorders. (Updated on 25-MAR-
CC 2003 to correct PN field.)
XX
SQ Sequence 1527 BP; 246 A; 514 C; 470 G; 297 T; 0 U; 0 Other;

Query Match   100.0%; Score 1527; DB 2; Length 1527;
Best Local Similarity   100.0%; Pred. No. 0;
Matches 1527; Conservative   0; Mismatches   0; Indels   0; Gaps   0;

```
Qy      1 ATGGACCACCTCGGGGCCGTCCCTCTGCCCCAGTCGCCTCCTTTGTCTCCTGCTCGCT 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 ATGGACCACCTCGGGGCCGTCCCTCTGCCCCAGTCGCCTCCTTTGTCTCCTGCTCGCT 60

Qy     61 GGGGCCGGCCGGGGCGCCCCCGGACCCCAAGTTCCCGGACCCCAAGTTGAGAGACAAAGCGGGC 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61 GGGGCCGGCCGGGGCGCCCCCGGACCCCAAGTTCCCGGACCCCAAGTTGAGAGACAAAGCGGGC 120

Qy    121 TTGCTGCGGCGGCCCGGGGGCCCGAAGAGTTCTGTGCTTCACCGAGCGGTTGGAGGACTTG 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121 TTGCTGCGGCGGCCCGGGGGCCCGAAGAGTTCTGTGCTTCACCGAGCGGTTGGAGGACTTG 180

Qy    181 GTGTGTTTCTGGGAGAAGCGGCGAGCGGCTGGGGTGGGCCCGGCCCCGGCCAACTACAGCTTCTCC 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    181 GTGTGTTTCTGGGAGAAGCGGCGAGCGGCTGGGGTGGGCCCGGCCCCGGCCAACTACAGCTTCTCC 240

Qy    241 TACCAGCTCGAGGATGAGCCAGTTGTGTCGCCTCCACGGCTCCACGGCTCGT 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    241 TACCAGCTCGAGGATGAGCCAGTTGTGTCGCCTCCACGGCTCCACGGCTCGT 300

Qy    301 GGTGCGGTGCGCTTCGTGTTCGCTGCCTACAGCCGACACGTGCGAGCTTGCGAGCTTGTGCCCCTA 360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    301 GGTGCGGTGCGCTTCGTGTTCGCTGCCTACAGCCGACACGTGCGAGCTTGCGAGCTTGTGCCCCTA 360

Qy    361 GAGTTGCGCGTCACAGCAGCCTCCGGCGCTCCGGCGCCGGATATCACCGTGTCATCCACATCAAT 420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    361 GAGTTGCGCGTCACAGCAGCCTCCGGCGCTCCGGCGCCGGATATCACCGTGTCATCCACATCAAT 420

Qy    421 GAAGTAGTGCTCCTAGAACGCGTGGGCGTTGGTGGCGGGCGGTTGGCTGACGAGAGCGGC 480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    421 GAAGTAGTGCTCCTAGAACGCGTGGGCGTTGGTGGCGGGCGGTTGGCTGACGAGAGCGGC 480

Qy    481 CACGTAGTGTTGCGCTGGCTGCTCCGGCCGCTCACGTCTCACATCCGCTAC 540
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    481 CACGTAGTGTTGCGCTGGCTGCTCCGGCCGCTCACGTCTCACATCCGCTAC 540

Qy    541 GAGGTGGACGTCTCGGCCGCGCAACGCGGCCAGGGAGAGCGTACAGAGGGTGGAGATCCTGGAG 600
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    541 GAGGTGGACGTCTCGGCCGCGCAACGCGGCCAGGGAGAGCGTACAGAGGGTGGAGATCCTGGAG 600

Qy    601 GGCCGCACCGAGTGTGTGCTGAGCAACGCGTACACCTGCGCGGGGCCGTACACCTTCGCGTC 660
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    601 GGCCGCACCGAGTGTGTGCTGAGCAACGCGTACACCTGCGCGGGGCCGTACACCTTCGCGTC 660

Qy    661 CGGCGCGTATGCGCTGAGCGAGCTTCGGCGGCCGCTTCGGAGGCGCGCCGGAGCCGTG 720
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    661 CGGCGCGTATGCGCTGAGCGAGCTTCGGCGGCCGCTTCGGAGGCGCGCCGGAGCCGTG 720

Qy    721 TCGCTGCTGACGGCCTAGCGACCTGGACCCCTCATCCTGACGCGCTCTCCCTCATCCTGTG 780
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    721 TCGCTGCTGACGGCCTAGCGACCTGGACCCCTCATCCTGACGCGCTCTCCCTCATCCTGTG 780

Qy    781 GTCATCCTGGTGCTGCTGACGCGCTCGACGGTGCTGCTGACGCGCTCGACGGCTCTGAAGCAG 840
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    781 GTCATCCTGGTGCTGCTGACGCGCTCGACGGTGCTGCTGACGCGCTCGACGGCTCTGAAGCAG 840

Qy    841 AAGATCTGGCCTGGCATCCCGAGCCCAGAGAGCGAGTTGAAGGCCTCTTCACCACCCAC 900
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    841 AAGATCTGGCCTGGCATCCCGAGCCCAGAGAGCGAGTTGAAGGCCTCTTCACCACCCAC 900

Qy    901 AAGGGTAACTTCACGTCTGGCGTGTACCAGATGATGGCTGCGTGTGGTGGGAGCCCTGC 960
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    901 AAGGGTAACTTCACGTCTGGCGTGTACCAGATGATGGCTGCGTGTGGTGGGAGCCCTGC 960

Qy    961 ACCCCCTTCACGGAGGACCCACCTGCTTCCCTGGAAGTCCTTCCCTGGAAGTCCTCAGAGCGTGCTGGGGG 1020
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    961 ACCCCCTTCACGGAGGACCCACCTGCTTCCCTGGAAGTCCTTCCCTGGAAGTCCTCAGAGCGTGCTGGGGG 1020

Qy   1021 ACGATGAGGCAGTGGAGCACCGGGGACAGATGATGAGGGCCCCCTGCTGGAGCCAGTGGGC 1080
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1021 ACGATGAGGCAGTGGAGCACCGGGGACAGATGATGAGGGCCCCCTGCTGGAGCCAGTGGGC 1080

Qy   1081 AGTGAGCATGCCCAGGATACCTATCTGGTCGGTGGACAAATGGTTGCTGCCCCGGAACCCG 1140
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1081 AGTGAGCATGCCCAGGATACCTATCTGGTCGGTGGACAAATGGTTGCTGCCCCGGAACCCG 1140

Qy   1141 CCCAGTGAGGACTTCCCAGGGCCTCCGTGTGGCAGTAGTGGCCATAGTGGACATGAGAAGGC 1200
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1141 CCCAGTGAGGACTTCCCAGGGCCTCCGTGTGGCAGTAGTGGCCATAGTGGACATGAGAAGGC 1200

Qy   1201 TCAGAAGCATCCTCCTGCTCATCGTTTGGCCTCGAAGCCCAGAGCCCAGCCAGAGGAGCCTCT 1260
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1201 TCAGAAGCATCCTCCTGCTCATCGTTTGGCCTCGAAGCCCAGAGCCCAGCCAGAGGAGCCTCT 1260

Qy   1261 GCTGCCAGCTTTGAGTACACTATCCTGACCCCCACCTAAAGTACCTGTGTACCTTGTGGTATCT 1320
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1261 GCTGCCAGCTTTGAGTACACTATCCTGACCCCCACCTAAAGTACCTGTGTACCTTGTGGTATCT 1320

Qy   1321 CTGTGCCCTGAGCTGCCCCACCCACCCCACCTAAAGTAGCCCAGGGGACTCAGGGGCCCAAGGGGGCTTA 1380
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1321 CTGTGCCCTGAGCTGCCCCACCCACCCCACCTAAAGTAGCCCAGGGGACTCAGGGGCCCAAGGGGGCTTA 1380

Qy   1381 GACTCTGGCATCTCAACTGACTAGGGACTCAGGGGCCCAAGGGGGCTTA 1440
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1381 GACTCTGGCATCTCAACTGACTAGGGACTCAGGGGCCCAAGGGGGCTTA 1440

Qy   1441 TCCGATGGCCCCTACTCCAACCCTTATGAGAACAGCAGCCTTATCCCGCTGAGCCCTCTG 1500
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1441 TCCGATGGCCCCTACTCCAACCCTTATGAGAACAGCAGCCTTATCCCGCTGAGCCCTCTG 1500

Qy   1501 CCCCCCAGCTATGTGCCTGCTCTTAG 1527
          ||||||||||||||||||||||||||
Db   1501 CCCCCCAGCTATGTGCCTGCTCTTAG 1527
```

```
PD   22-JUL-2004.
XX
PF   15-OCT-2003; 2003WO-US029126.
XX
PR   18-OCT-2002; 2002US-0418988P.
XX
PA   (GETH ) GENENTECH INC.
PA   (WUTD/) WU T D.
PA   (ZHOU/) ZHOU Y.
XX
PI   Wu TD, Zhou Y;
XX
DR   WPI; 2004-534300/51.
XX
PT   New nucleic acid molecule and encoded polypeptide, for diagnosing,
PT   preventing or treating cell proliferative disorders such as cancer.
XX
PS   Claim 1; SEQ ID NO 360; 5504pp; English.
XX
CC   The present invention describes an isolated tumour-associated antigenic
CC   target (TAT) nucleic acid comprising: (a) any of 4622 nucleotide
CC   sequences (see SEQ ID NO:1 to 4622); (b) the full-length coding region of
CC   (a); (c) the complement of (a) or (b); (d) a sequence that has 80%
CC   sequence identity to (a)-(c); or (e) a sequence that hybridises to (a)-
CC   (c). Also described: (1) an expression vector comprising the above
CC   nucleic acid; (2) a host cell comprising the above expression vector; (3)
CC   a process for producing a polypeptide; (4) an isolated polypeptide
CC   comprising: (a) an amino acid sequence encoded by any of the above
CC   nucleotide sequences; (b) an amino acid sequence encoded by the full-
CC   length coding region of the above nucleotide sequences; or (c) a sequence
CC   having at least 80% identical to (a) or (b); (5) a chimeric polypeptide
CC   comprising the above polypeptide fused to a heterologous polypeptide; (6)
CC   an isolated antibody that binds to the above polypeptide; (7) a process
CC   for producing the antibody; (8) an isolated oligopeptide that binds to
CC   the above polypeptide; (9) a tumour-associated antigenic target (TAT)
CC   binding organic molecule that binds to the above polypeptide; (10) a
CC   composition of matter comprising the above (chimeric) polypeptide,
CC   antibody, oligopeptide or TAT binding organic molecule, in combination
CC   with a carrier; (11) an article of manufacture comprising a container and
CC   the composition of matter contained within the container; (12) methods of
CC   inhibiting the growth of a cell that expresses the above protein, where
CC   the growth of the cell is at least in part dependent upon a growth
CC   potentiating effect of the above protein; (13) a method of
CC   therapeutically treating a mammal having a cancerous tumour comprising
CC   cells that express the above protein; (14) a method of determining the
CC   presence of a protein in a sample suspected of containing the protein
CC   described above; (15) methods of diagnosing the presence of a tumour in a
CC   mammal; (16) a method for treating or preventing a cell proliferative
CC   disorder associated with increased expression or activity of the above
CC   protein; and (17) a method of binding an antibody, oligopeptide or
CC   organic molecule to a cell that expresses the protein described above.
CC   The TAT sequences have cytostatic activities, and can be used in gene
CC   therapy. The composition and methods are useful for diagnosing,
CC   preventing or treating cancer. The composition is also used for preparing
CC   a medicament for the therapeutic treatment or diagnostic detection of a
CC   cell proliferative disorder or cancer. The present sequence represents a
CC   human TAT cDNA sequence from the present invention.
XX
SQ   Sequence 1848 BP; 313 A; 593 C; 577 G; 365 T; 0 U; 0 Other;

Query Match      100.0%;   Score 1527;  DB 13;   Length 1848;
Best Local Similarity   100.0%;   Pred. No: 0;
Matches 1527   Conservative   0;   Mismatches   0;   Indels   0;   Gaps   0;

Qy        1 ATGGACCACCTCGGGGCGTCCCTCTGGCCCCCAGGTCGGCTCCCTTTGTCTCCTGCTCGCT 60
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      137 ATGGACCACCTCGGGGCGTCCCTCTGGCCCCCAGGTCGGCTCCCTTTGTCTCCTGCTCGCT 196

Qy       61 GGGGCCGCGCCTGGGCGCCCCCCGCCTAACCTCCCGGACCCCAGTTCGAGAGCAAAGCGCC 120
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      197 GGGGCCGCGCCTGGGCGCCCCCCGCCTAACCTCCCGGACCCCAGTTCGAGAGCAAAGCGCC 256

Qy      121 TTGCTGGCGGCCCGGCCCGGGGCCCGAAGAGCTTCTGTGCTTCACCGAGCGGTTGGAGGACTTG 180
```

(Note: the above alignment blocks continue in the upper portion of the page, shown rotated, as follows:)

```
Db      257 TTGCTGGCGGCCCGGGGCCCGAAGAGCTTCTGTGCTTCACCGAGCGGTTGGAGGACTTG 316

Qy      181 GTGTGTTTCTGGAGGAAGGGGAGCGGCGAGCGGTTGGCCGGGCAACTACAGCTTCTCC 240
Db      317 GTGTGTTTCTGGAGGAAGGGGAGCGGCGAGCGGTTGGCCGGGCAACTACAGCTTCTCC 376

Qy      241 TACCAGCTCGAGGATGAGCCATGAGCGTGTCGCTGCCTACAGCCGACAGTCGAGCTTCGTGCCCGGCTCGT 300
Db      377 TACCAGCTCGAGGATGAGCCATGAGCGTGTCGCTGCCTACAGCCGACAGTCGAGCTTCGTGCCCGGCTCGT 436

Qy      301 GGTGCGGTGCGCTTCGCTGGTGTTCGCTGCCTACAGCCGACACAGTCGAGCTTCGTGTGCCCCTA 360
Db      437 GGTGCGGTGCGCTTCGCTGGTGTTCGCTGCCTACAGCCGACACAGTCGAGCTTCGTGTGCCCCTA 496

Qy      361 GAGTTGCGCGTCACAGCAGCCTCCGGCGCTCCGGCGCTCCGGATATCACCGTGTCATCCACATCAAT 420
Db      497 GAGTTGCGCGTCACAGCAGCCTCCGGCGCTCCGGCGCTCCGGATATCACCGTGTCATCCACATCAAT 556

Qy      421 GAAGTAGTGCTTCCTAGACGACGCCCCCGTGGGGCTGGTCGGCGCTGGTCGGCTGGTTGGCTGCTGGAGGAGCGGGC 480
Db      557 GAAGTAGTGCTTCCTAGACGACGCCCCCGTGGGGCTGGTCGGCGCTGGTCGGCTGGTTGGCTGCTGGAGGAGCGGGC 616

Qy      481 CACGTAGTGTTGCGCTGGCTGGCTCCGAGACACCCATGAGCGTCTCACATCCGCTAC 540
Db      617 CACGTAGTGTTGCGCTGGCTGGCTCCGAGACACCCATGAGCGTCTCACATCCGCTAC 676

Qy      541 GAGGTGGAGTCTCGGCGGCCGCAACGGCCAGGGAGCGTACAGAGGGTGGAGATCCTGGAG 600
Db      677 GAGGTGGAGTCTCGGCGGCCGCAACGGCCAGGGAGCGTACAGAGGGTGGAGATCCTGGAG 736

Qy      601 GGCGCCACCGAGTGTGTGCTGAGCAACCTGCGGGGCCGGACGCGCTACAACCTTCGCCGTC 660
Db      737 GGCGCCACCGAGTGTGTGCTGAGCAACCTGCGGGGCCGGACGCGCTACAACCTTCGCCGTC 796

Qy      661 CGGCGGCGTATGGCTGAGCCGAGCTTCGGCGGCTTCTGAGCGCGGCGCCGGTCGGAGCCGTGTG 720
Db      797 CGGCGGCGTATGGCTGAGCCGAGCTTCGGCGGCTTCTGAGCGCGGCGCCGGTCGGAGCCGTGTG 856

Qy      721 TCGCTGCTGACGCCTAGGACCTGGACCCCCCATCCTGACGCGCTCCCTCGTG 780
Db      857 TCGCTGCTGACGCCTGCTGACGCCTGGACCCCCCATCCTGACGCGCTCCCTCGTG 916

Qy      781 GTCATCCTGCTGCTGACCGTGCTGCTGCTGACCGTCTCCCACCGCCGGGCTCTGAAGCAG 840
Db      917 GTCATCCTGCTGCTGACCGTGCTGCTGCTGACCGTCTCCCACCGCCGGGCTCTGAAGCAG 976

Qy      841 AAGATCTGGCCGGATCCCGAGCCCAGAGAGCGAGTTTGAAGGCCTCTTCACCACCACCAC 900
Db      977 AAGATCTGGCCGGATCCCGAGCCCAGAGAGCGAGTTTGAAGGCCTCTTCACCACCACCAC 1036

Qy      901 AAGGTAACTTCAGCTGTGGCTGTACCAGAATGATGGCTGCCTGTGGTGGTGAGCCCCTGC 960
Db     1037 AAGGTAACTTCAGCTGTGGCTGTACCAGAATGATGGCTGCCTGTGGTGGTGAGCCCCTGC 1096

Qy      961 ACCCCCTTCACGGAGGAGGACCACCTGCTTCCTCCTGAAGTCCTCTCCAGAGAGCCTGCTGGGGGG 1020
Db     1097 ACCCCCTTCACGGAGGAGGACCACCTGCTTCCTCCTGAAGTCCTCTCCAGAGAGCCTGCTGGGGGG 1156

Qy     1021 ACGATGCAGCAGTGGAGCCAGGACAGATGAGGGCCCCTGCTGAGCCAGTGGGC 1080
Db     1157 ACGATGCAGCAGTGGAGCCAGGACAGATGAGGGCCCCTGCTGAGCCAGTGGGC 1216

Qy     1081 AGTGAGCAGTGCCCAGGATACCTATCTGGTGCTGACAAATGGTTGCTGCCCCGGAACCCG 1140
Db     1217 AGTGAGCAGTGCCCAGGATACCTATCTGGTGCTGACAAATGGTTGCTGCCCCGGAACCCG 1276

Qy     1141 CCCAGTGAGGACCTTCCCAGGGCCTGGTGCTGTGGACATAGTGGCCATGGATGAAGGC 1200
Db     1277 CCCAGTGAGGACCTTCCCAGGGCCTGGTGCTGTGGACATAGTGGCCATGGATGAAGGC 1336

Qy     1201 TCAGAAGCATCCTCCTGCTTCATCTGCTTTGCCTGAAGCCGAGCGGTTGGAGGACTCT 1260
```

CC   benefits from modulating the balance of regulatory T cell function
CC   relative to effector T cell function, or vice versa, in a subject. The
CC   method involves administering an agent that modulates the expression or
CC   activity of a molecule selected from PTGER2 and TGFbeta1, or Jagged-1,
CC   GPR-32, CD83, CD84, CD89, serotonin R, BY55, serotonin R2C, GPR63,
CC   histamine R-H4, GPR58, EPO-R, PSG-1, PSG-3, PSG-6, PDE-4d, and PI-
CC   3-related kinase to the subject, such that treatment occurs. The methods
CC   are useful for diagnosing, preventing or treating conditions
CC   characterized by a too-vigorous or weak effector T cell or regulatory T
CC   cell response to antigens associated with the condition, such as in an
CC   allergic response, an autoimmune disorder, a viral infection, a microbial
CC   infection, a parasitic infection or a tumour. The present sequence
CC   represents a DNA encoding a human erythropoietin receptor (EPOR),
CC   preferentially expressed in regulatory T cells.
XX
SQ   Sequence 1849 BP; 313 A; 593 C; 578 G; 365 T; 0 U; 0 Other;

Query Match   100.0%; Score 1527; DB 12; Length 1849;
Best Local Similarity   100.0%; Pred. No. 0;
Matches 1527; Conservative   0; Mismatches   0; Indels   0; Gaps   0;

```
Qy         1 ATGGACCACCTCGGGGCGTCCCTCCTGCCCCAGGTCGGCTCCCTTGTCTCCTGCTCGCT 60
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       137 ATGGACCACCTCGGGGCGTCCCTCCTGCCCCAGGTCGGCTCCCTTGTCTCCTGCTCGCT 196

Qy        61 GGGGCCGCTCGGGGCGCCCGGGGCCCCCAGTTCGAGAGCAAAGCGGCC 120
             ||||||||||||||||||||||||||||||||||||||||||||||||
Db       197 GGGGCCGCTCGGGGCGCCCGGGGCCCCCAGTTCGAGAGCAAAGCGGCC 256

Qy       121 TTGCTGGCGGGGCCCGGGGGCGCCCGAAGAGCTTCTGTGCTTCACCGAGCCGGTTGGAGACTTG 180
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       257 TTGCTGGCGGGGCCCGGGGGCGCCCGAAGAGCTTCTGTGCTTCACCGAGCCGGTTGGAGACTTG 316

Qy       181 GTGTGTTTCTGGAGAAGCGGCGAGCGGCTGGGGTGGGCCCGGGCAACTACAGCTTCTCC 240
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       317 GTGTGTTTCTGGAGAAGCGGCGAGCGGCTGGGGTGGGCCCGGGCAACTACAGCTTCTCC 376

Qy       241 TACCAGCTCGAAGATGAGCAAGCCATGGCCGTGTGTCGCTCGCCGCCCACCAGGCTCCCACGGCTCGT 300
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       377 TACCAGCTCGAAGATGAGCAAGCCATGGCCGTGTGTCGCTCGCCGCCCACCAGGCTCCCACGGCTCGT 436

Qy       301 GGTGCGGTGCGCTTCGTGTTCGCTGCTCAGCGCCGACACGTCGAGCTTCGTGTTCGTGCCCCCTA 360
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       437 GGTGCGGTGCGCTTCGTGTTCGCTGCTCAGCGCCGACACGTCGAGCTTCGTGTTCGTGCCCCCTA 496

Qy       361 GAGTTGCGCGTTCTAGACGCCTCCGGCGCTCCGGCGGGCTGGTGGGGCGGTTCATCGTCATCACCATCAAT 420
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       497 GAGTTGCGCGTTCTAGACGCCTCCGGCGCTCCGGCGGGCTGGTGGGGCGGTTCATCGTCATCACCATCAAT 556

Qy       421 GAAGTAGTGCTCCTAGACGCCCCGTGGTGGGCGCGGTTGGTGGCGGAGAGCGCGC 480
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       557 GAAGTAGTGCTCCTAGACGCCCCGTGGTGGGCGCGGTTGGTGGCGGAGAGCGCGC 616

Qy       481 CACGTAGTGTTGCGCTGCTCCTGAGACGTCTCACCATGAGTCTCACATCCGCTAC 540
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       617 CACGTAGTGTTGCGCTGCTCCTGAGACGTCTCACCATGAGTCTCACATCCGCTAC 676

Qy       541 GAGGTGGACGTCTCGGCGCCGGCAACGGCGCAGGGAGCGCAGGAGGGGTGAGACGGGTGAGATCCTGGAG 600
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       677 GAGGTGGACGTCTCGGCGCCGGCAACGGCGCAGGGAGCGCAGGAGGGGTGAGACGGGTGAGATCCTGGAG 736

Qy       601 GGCGGCCACCGAGTGTGCTGAGCAACGACCTTGTCGGCGGGGCCGGCACCACCTTCGCGCGTC 660
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       737 GGCGGCCACCGAGTGTGCTGAGCAACGACCTTGTCGGCGGGGCCGGCACCACCTTCGCGCGTC 796

Qy       661 CGGCGGCGTATGCGGCTTCGGAGCCTGTGGAGGCCGTGTG 720
             ||||||||||||||||||||||||||||||||||||||||
Db       797 CGGCGGCGTATGCGGCTTCGGAGCCTGTGGAGGCCGTGTG 856

Qy       721 TCGCTGCTGACGGCTCAGCGACCTGACGCCCCTCATCCTGACGTCTTCTTCCCTGACGTCTCATCCTCGTG 780
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       857 TCGCTGCTGACGGCTCAGCGACCTGACGCCCCTCATCCTGACGTCTTCTTCCCTGACGTCTCATCCTCGTG 916
```

```
Db      1337 TCAGAAGCATCCTCCTGCTCATCTGCTTTGGCCTCGAAGCCCAGCGAGGGAGCCTCT 1396
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy      1261 GCTGCCAGCTTTGAGTACACTATCCTGGACCCCAGCTCCCAGCTCTTGCTCCATGGACA 1320
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1397 GCTGCCAGCTTTGAGTACACTATCCTGGACCCCAGCTCCCAGCTCTTGCTCCATGGACA 1456

Qy      1321 CTGTGCCCTGAGCTGCGCCCCTACCCCACCCCACCCCACCCTAAAGTACCTGTACCTGTGGTATCT 1380
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1457 CTGTGCCCTGAGCTGCGCCCCTACCCCACCCCACCCCACCCTAAAGTACCTGTACCTGTGGTATCT 1516

Qy      1381 GACTCTGGCATCTCAACTGACTACAGCTCAGGGGACTCCCAGGGAGCCCAAGGGGCTTA 1440
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1517 GACTCTGGCATCTCAACTGACTACAGCTCAGGGGACTCCCAGGGAGCCCAAGGGGCTTA 1576

Qy      1441 TCCGATGGCCCCTACTCCAACCCTTATGAGAACAGCCTTATCCCAGCCGCTGAGCCTCTG 1500
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1577 TCCGATGGCCCCTACTCCAACCCTTATGAGAACAGCCTTATCCCAGCCGCTGAGCCTCTG 1636

Qy      1501 CCCCCAGCTATGTGGCTTGCTTCTTAG 1527
             |||||||||||||||||||||||||||
Db      1637 CCCCCAGCTATGTGGCTTGCTTCTTAG 1663
```

RESULT 3
ADO05723
ID   ADO05723 standard; DNA; 1849 BP.
XX
AC   ADO05723;
XX
DT   15-JUL-2004 (first entry)
XX
DE   Human erythropoietin receptor (EPOR) encoding DNA.
XX
KW   T cell; antiallergic; immunosuppressive; virucide; antibacterial;
KW   antiparasitic; cytostatic; gene therapy; human; gene; ds;
KW   erythropoietin receptor; EPOR.
XX
OS   Homo sapiens.
XX
FH   Key   Location/Qualifiers
FH
FT   CDS   137..1663
FT   /tag= a
FT   /product= "erythropoietin receptor"
XX
PN   WO2004032867-A2.
XX
PD   22-APR-2004.
XX
PF   09-OCT-2003; 2003WO-US032065.
XX
PR   09-OCT-2002; 2002US-0417102P.
PR   09-OCT-2002; 2002US-0417103P.
PR   09-OCT-2002; 2002US-0417243P.
PR   09-OCT-2002; 2002US-0419575P.
PR   18-OCT-2002; 2002US-0424777P.
PR   08-NOV-2002; 2002US-0424881P.
PR   08-NOV-2002; 2002US-0424881P.
XX
PA   (TOLE-) TOLERRX INC.
XX
PI   Rao P, Szymanska G;
XX
DR   WPI; 2004-340801/31.
DR   P-PSDB; ADO05724.
DR   GENBANK; 4557561.
XX
PT   Treating a condition that benefits from modulating regulatory or effector
PT   T cell function comprises administering an agent that modulates the
PT   expression or activity of a gene or polypeptide (e.g. PTGER2, TGFbeta1,
PT   Jagged-1 or GPR-32).
XX
PS   Example; SEQ ID NO 23; 161pp; English.
XX
CC   The invention relates to treating a subject having a condition that

XX
PR   24-APR-2002; 2002US-00131831.
PR   20-DEC-2002; 2002US-0325899.
XX
PA   (EXPR-) EXPRESSION DIAGNOSTICS INC.
XX
PI   Wohlgemuth J, Fry K, Woodward R, Ly N, Prentice J, Morris M;
PI   Rosenberg S;
XX
DR   WPI; 2004-400724/37.
XX
PT   Diagnosing or monitoring transplant rejection, e.g. heart, kidney, liver,
PT   pancreas, pancreatic islet, lung, bone marrow or stem cell transplant
PT   rejection, in an individual, comprises detecting the expression level of
PT   the genes.
XX
PS   Claim 80; SEQ ID NO 360; 1762pp; English.
XX
CC   The present invention relates to diagnosing or monitoring transplant
CC   rejection, e.g. cardiac or kidney transplant rejection, in an individual
CC   comprises detecting the expression level of one or more genes. The
CC   methods, system and kits are useful in diagnosing or monitoring
CC   transplant rejection, e.g. heart, kidney, liver, pancreas, pancreatic
CC   islet, lung, bone marrow or stem cell transplant rejection,
CC   xenotransplant rejection or mechanical organ replacement rejection, in an
CC   individual. The method is also useful in assessing the immune status of
CC   an individual. The methods are also useful in diagnosing and monitoring
CC   diseases that involve the immune system, e.g. rheumatoid arthritis,
CC   lupus, inflammatory bowel diseases, multiple sclerosis, HIV/AIDS or
CC   viral, bacterial or fungal infection. The present sequence represents a
CC   mRNA reference sequence for a 50 mer oligonucleotide marker for diagnosis
CC   and monitoring of allograft rejection and other disorders.
XX
SQ   Sequence 1849 BP; 313 A; 593 C; 578 G; 365 T; 0 U; 0 Other;

Query Match   100.0%;   Score 1527;   DB 12;   Length 1849;
Best Local Similarity 100.0%;   Pred. No. 0;
Matches 1527;   Conservative   0;   Mismatches   0;   Indels   0;   Gaps   0;

Qy   1 ATGGACCACCTCGGGGCGTCCCCTCTGCCCCAGGTGGCTCGCTCCCTTGTCTCCTGCTCGCT   60
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   137 ATGGACCACCTCGGGGCGTCCCCTCTGCCCCAGGTGGCTCGCTCCCTTGTCTCCTGCTCGCT 196

Qy   61 GGGGCCCGCCTGGGGCCCCGGGGGCCCGCCTGGGGCCCCGGGGGCCCCAGTTCGAGAGCAAAGCGGCC 120
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   197 GGGGCCCGCCTGGGGCCCCGGGGGCCCGCCTGGGGCCCCGGGGGCCCCAGTTCGAGAGCAAAGCGGCC 256

Qy   121 TTGCTGGCGGCCCGGGGCCCGGGGCCCCGAAGAGCTTCTGCTTCACCGAGCGGTTGGAGGACTTG 180
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   257 TTGCTGGCGGCCCGGGGCCCGGGGCCCCGAAGAGCTTCTGCTTCACCGAGCGGTTGGAGGACTTG 316

Qy   181 GTGTGTTTCTGGAGAAGCGGCGAGCGGCGGAGCGGTGGGCCCGGACCACAGCTTCTCC 240
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   317 GTGTGTTTCTGGAGAAGCGGCGAGCGGCGGAGCGGTGGGCCCGGACCACAGCTTCTCC 376

Qy   241 TACCAGCTCGAGGATGAGCGCCATGAGCGTGTGCCTACAGCGCCACGGCTCGT 300
         |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   377 TACCAGCTCGAGGATGAGCGCCATGAGCGTGTGCCTACAGCGCCACGGCTCGT 436

Qy   301 GGTGCGGTGCCGTTCGTGTGTGGCGTTCGTCGCTCCACAGCGCACGTCGAGCTTCGTGCCCCTA 360
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   437 GGTGCGGTGCCGTTCGTGTGTGGCGTTCGTCGCTCCACAGCGCACGTCGAGCTTCGTGCCCCTA 496

Qy   361 GAGTTGCGGTCCACAGCGGCCTCCGGATATCATCCGTGTCATCCACATCAAT 420
         |||||||||||||||||||||||||||||||||||||||||||||||||||
Db   497 GAGTTGCGGTCCACAGCGGCCTCCGGATATCATCCGTGTCATCCACATCAAT 556

Qy   421 GAAGTAGTAGTCCTAGACGCCCCGTCGGGGCGTCGGGGCGGTTGGCGGGACTTGGCTGGCTGGAGAGCGGC 480
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   557 GAAGTAGTAGTCCTAGACGCCCCGTCGGGGCGTCGGGGCGGTTGGCGGGACTTGGCTGGCTGGAGAGCGGC 616

Qy   481 CACGTAGTTGCGCTGCCTCCCGCCCGAGACACCCATGAGCGTCTCACATCCGCTAC 540
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||

Qy   781 GTCATCCTGGTGCTGCTGACCGTGCTGCGCGTGCTCTCCCACCGCGGCTCTGAAGCAG 840
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   917 GTCATCCTGGTGCTGCTGACCGTGCTGCGCGTGCTCTCCCACCGCGGCTCTGAAGCAG 976

Qy   841 AAGATCTGGCCTGGCATCCCGAGCCCAGAGAGCGAGTTTGAAGGCCTCTTCACCACCCAC 900
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   977 AAGATCTGGCCTGGCATCCCGAGCCCAGAGAGCGAGTTTGAAGGCCTCTTCACCACCCAC 1036

Qy   901 AAGGGTAACTTCCAGCTGTGGCGTGTACCAGATGATGATGCTGCCTGTGTGGAGCCCCTGC 960
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1037 AAGGGTAACTTCCAGCTGTGGCGTGTACCAGATGATGATGCTGCCTGTGTGGAGCCCCTGC 1096

Qy   961 ACCCCTTCACGGAGGAGACCACCACCTGCTTCCCTGGAAGTCCTTCAGAGCGCTGCTGGGGG 1020
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1097 ACCCCTTCACGGAGGAGACCACCACCTGCTTCCCTGGAAGTCCTTCAGAGCGCTGCTGGGGG 1156

Qy 1021 ACGATGAGGCAGTGGGAGCCGGGACAGATGATGAGGGCCCCCTGCTGGAGCCAGTGGGC 1080
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1157 ACGATGAGGCAGTGGGAGCCGGGACAGATGATGAGGGCCCCCTGCTGGAGCCAGTGGGC 1216

Qy 1081 AGTGAGCATGCCCAGGATACCTATCTGTGCTGGACAAATGGTTGCTGCCCCGGAACCCG 1140
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1217 AGTGAGCATGCCCAGGATACCTATCTGTGCTGGACAAATGGTTGCTGCCCCGGAACCCG 1276

Qy 1141 CCCAGTGAGGACCTCCCAGGGCCTGGTGGCACATAGTGGCCATGGATGAAGGC 1200
         ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1277 CCCAGTGAGGACCTCCCAGGGCCTGGTGGCACATAGTGGCCATGGATGAAGGC 1336

Qy 1201 TCAGAAGCATCCTCCTGCTCATCTGCTTTGGCCTGAAGCCCAGAGAGGAGCCTCT 1260
         |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1337 TCAGAAGCATCCTCCTGCTCATCTGCTTTGGCCTGAAGCCCAGAGAGGAGCCTCT 1396

Qy 1261 GCTGCCAGCTTTGAGTACATATCCTGACCCCCAGCTCTTGCGTCCCATGACA 1320
         |||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1397 GCTGCCAGCTTTGAGTACATATCCTGACCCCCAGCTCTTGCGTCCCATGACA 1456

Qy 1321 CTGTGCCCTGAGCTGCCCCCTACCCCACCCCACCTAAAGTACCTGTACCTTGTGGTATCT 1380
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1457 CTGTGCCCTGAGCTGCCCCCTACCCCACCCCACCTAAAGTACCTGTACCTTGTGGTATCT 1516

Qy 1381 GACTCTGCATCTCAACTGACTACAGCTCAGGGAGCTCCCAGGGAGCCCAAGGGGGCTTA 1440
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1517 GACTCTGCATCTCAACTGACTACAGCTCAGGGAGCTCCCAGGGAGCCCAAGGGGGCTTA 1576

Qy 1441 TCCGATGGCCCCCTACTCCAACCCTTATGAGACAACAGCCTTATCCCAGCCGCTGAGCCTCTG 1500
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1577 TCCGATGGCCCCCTACTCCAACCCTTATGAGACAACAGCCTTATCCCAGCCGCTGAGCCTCTG 1636

Qy 1501 CCCCCAGCTATGTGGCTTGCTCTTAG 1527
         ||||||||||||||||||
Db 1637 CCCCCAGCTATGTGGCTTGCTCTTAG 1663

RESULT 4
ADP10351
ID   ADP10351 standard; DNA; 1849 BP.
XX
AC   ADP10351;
XX
DT   12-AUG-2004 (first entry)
XX
DE   Reference mRNA sequences for marker probe #28.
XX
KW   transplant rejection; immune system; rheumatoid arthritis; lupus;
KW   inflammatory bowel disease; multiple sclerosis; HIV; AIDS; ss.
XX
OS   Homo sapiens.
XX
PN   WO200404234G-A2.
XX
PD   21-MAY-2004.
XX
PF   24-APR-2003; 2003WO-US012946.

ID	ACN40465 standard; cDNA; 1849 BP.
XX	
AC	ACN40465;
XX	
DT	18-NOV-2004 (first entry)
XX	
DE	Tumour-associated antigenic target (TAT) cDNA DNA218271, SEQ ID NO:5258.
XX	
KW	Tumour-associated antigenic target; TAT; human; overexpression; cancer;
KW	tumour; diagnosis; cell proliferative disorder; breast cancer;
KW	colorectal cancer; lung cancer; ovarian cancer; liver cancer;
KW	central nervous system cancer; bladder cancer; pancreatic cancer;
KW	cervical cancer; melanoma; leukaemia; hybridisation probe;
KW	chromosome identification; chromosome mapping; gene mapping;
KW	gene therapy; cytostatic; gene; ss.
XX	
OS	Homo sapiens.
XX	
PN	WO200403061S-A2.
XX	
PD	15-APR-2004.
XX	
PF	29-SEP-2003; 2003WO-US028547.
XX	
PR	02-OCT-2002; 2002US-0414971P.
XX	
PA	(GETH) GENENTECH INC.
XX	
PI	Wu TD, Zhang Z, Zhou Y;
XX	
DR	WPI; 2004-347921/32.
DR	P-PSDB; ABM82037.
XX	
PT	New tumour-associated antigenic target polypeptides and nucleic acids,
PT	useful in preparing a medicament for treating or detecting a
PT	proliferative disorder, e.g. breast, lung, colorectal, ovarian or
PT	prostate cancer or tumor.
XX	
PS	Claim 1; SEQ ID NO 5258; 7273pp; English.
XX	
CC	The invention relates to human tumour-associated antigenic target (TAT)
CC	polypeptides, and their related nucleic acids. The TAT polypeptides are
CC	overexpressed in cancer tissues compared to normal tissues, and may thus
CC	serve as effective targets for the diagnosis and treatment of cancer in
CC	mammal. The invention also relates to nucleic acid and polypeptide
CC	sequences at least 80% identical to the TAT nucleic acids and
CC	polypeptides; expression vectors and host cells comprising a TAT nucleic
CC	acid; an antibody specific for a TAT polypeptide; a peptide or organic
CC	molecule which binds to a TAT polypeptide; fusion proteins comprising a
CC	TAT polypeptide; and methods and compositions for the treatment or
CC	diagnosis of cancer in mammals. TAT polypeptides, nucleic acids,
CC	antibodies, antagonists, binding molecules and compositions are useful
CC	for diagnosing or treating a cell proliferative disorder associated with
CC	increased TAT expression, particularly cancers such as breast cancer,
CC	colorectal cancer, lung cancer, ovarian cancer, liver cancer, bladder
CC	cancer, pancreatic cancer, cervical cancer, cancers of the central
CC	nervous system, melanoma and leukaemia. TAT nucleic acids may further be
CC	used as hybridisation probes, in chromosome and gene mapping, in
CC	chromosome identification and in gene therapy. The present sequence
CC	represents a TAT nucleic acid of the invention
XX	
SQ	Sequence 1849 BP; 313 A; 593 C; 578 G; 365 T; 0 U; 0 Other;

Query Match   100.0%;   Score 1527;   DB 13;   Length 1849;
Best Local Similarity   100.0%;   Pred. No. 0;
Matches 1527;   Conservative   0;   Mismatches   0;   Indels   0;   Gaps   0;

```
Qy          1 ATGGACCACCTTCGGGGCGGGGTCCCTGCCCAGGTCGGCTCCCTTTGTCTCCTTGCTCGCT 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        137 ATGGACCACCTTCGGGGCGGGGTCCCTGCCCAGGTCGGCTCCCTTTGTCTCCTTGCTCGCT 196

Qy         61 GGGGCCGCTCGGGCCTGGGCGCGCCCCGGCTAACCTCCCGGACCCCAAGTTCGAGAGCAAAGCGGCC 120
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

Db   197 GGGGCCGCCTGGGCGCCCCGGCGCCCTAACCTCCCGGACCCCAGTTCGAGAGCAAAGCGGCC 256
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   121 TTGCTGGCGGCCCGGGGGCCCGAAGAGCTTCTGTGCTTCACCGAGCCGGTTGAGGGACTTG 180
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   257 TTGCTGGCGGCCCGGGGGCCCGAAGAGCTTCTGTGCTTCACCGAGCCGGTTGAGGGACTTG 316

Qy   181 GTGTGTTTCTGGGAGGAAGCGGCGAGCGGCTGGGCCCGGCAACTACAGCTTCTCC 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   317 GTGTGTTTCTGGGAGGAAGCGGCGAGCGGCTGGGCCCGGCAACTACAGCTTCTCC 376

Qy   241 TACCAGCTGAGGATGAGCCATGGAAGCTGTGTCGCCTGACCAGGTCTCCACGGCTCGT 300
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   377 TACCAGCTGAGGATGAGCCATGGAAGCTGTGTCGCCTGCCACCAGGTCTCCACGGCTCGT 436

Qy   301 GGTGCGGTGCGCTTCTGTGTTCGCTGCCTACAGCGTCGAGCTTCGTGCCCCTA 360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   437 GGTGCGGTGCGCTTCTGTGTTCGCTGCCTACAGCCGACAGTCGAGCTTCGTGCCCCTA 496

Qy   361 GAGTTGCGCGTCACAGCAGCCTCCGCGGCTCCGCGATATCACCGTGTCATCCACATCAAT 420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   497 GAGTTGCGCGTCACAGCAGCCTCCGCGGCTCCGCGATATCACCGTGTCATCCACATCAAT 556

Qy   421 GAAGTAGTGCTCCTAGACGCCCCCGTGGGCGTGGTGCGCGGTTGGCTGACGAGAGCGGC 480
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   557 GAAGTAGTGCTCCTAGACGCCCCCGTGGGGCTGGTGCGCGGTTGGCTGACGAGAGCGGGC 616

Qy   481 CACGTAGTGTTGCGCGTGGCTGGCTCCCGCGCCCTGAGACACACCCATGACGTCTCACATCCGCTAC 540
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   617 CACGTAGTGTTGCGCGTGGCTGGCTCCCGCGCCCTGAGACACACCCATGACGTCTCACATCCGCTAC 676

Qy   541 GAGGTGGACGTTCGCGGCCGGCAACGCGGCCAGGGAGCCTACAGAGGGTGAGATCCTGGAG 600
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   677 GAGGTGGACGTTCGCGGCCGGCAACGCGGCCAGGGAGCCTACAGAGGGTGAGATCCTGGAG 736

Qy   601 GGCCGCACCGAGTGTGTGCTGCTGAGCAACTGCGGGGGCCGGACGCGCTACACCTTCGCCGTC 660
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   737 GGCCGCACCGAGTGTGTGCTGCTGAGCAACTGCGGGGGCCGGACGCGCTACACCTTCGCCGTC 796

Qy   661 CGGCGGCGGTATGGCGTGCTGAGCGGCAGCTTCGGCCGGCTTCGGGAGGCGCCTGTG 720
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   797 CGCGGCGGTATGGCGTGCTGAGCGGCAGCTTCGGCCGGCTTCGGGAGGCGCCTGTG 856

Qy   721 TCGCTGCTGCTGAGCGCCTAGCGACCTGGACCCCCTCGACGCTCTCCCTCATCCTGTG 780
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   857 TCGCTGCTGCTGAGCGCCTAGCGACCTGGACCCCCTCGACGCTCTCCCTCATCCTGTG 916

Qy   781 GTCATCCTGGTGCTGCTGACCCGTCGCTGCTCTCCACCGCCGGGCTCTGAAGCAG 840
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   917 GTCATCCTGGTGCTGCTGACCCGTCGCTGCTCTCCACCGCCGGGCTCTGAAGCAG 976

Qy   841 AAGATCTGGCCTGGCATCCCAGACCCAGACGAGTTTGAAGGCCTCTTCACCACCCAC 900
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   977 AAGATCTGGCCTGGCATCCCAGACCCAGACGAGTTTGAAGGCCTCTTCACCACCCAC 1036

Qy   901 AAGGGTAACTTCAGCTGTGGCTGTACCAGAATGATGGCTGCCTGTGTGGAGCCCTGC 960
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1037 AAGGGTAACTTCAGCTGTGGCTGTACCAGAATGATGGCTGCCTGTGTGGGTGGAGCCCTGC 1096

Qy   961 ACCCCCTTCACGGGAGCAGTGGAGCTGCTCCTCAGAGTCCTCTCAGAGCCGCTGCTGGGGG 1020
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1097 ACCCCCTTCACGGGAGCAGTGGAGCTGCTCCTCAGAGTCCTCTCAGAGCCGCTGCTGGGGG 1156

Qy 1021 ACGATGCAGGCAGCAGTGAGCAGGGGACAGGCCGGGGACAGATGATGAGGGGCCCCTGCTGTGGGGC 1080
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1157 ACGATGCAGGCAGCAGTGAGCAGGGGACAGGCCGGGGACAGATGATGAGGGGCCCCTGCTGTGGGGC 1216

Qy 1081 AGTGAGCAGCATGCCAGATACCTATCTGTGCTGGACAAATGGTTGCTGCCCGGGAACCCG 1140
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1217 AGTGAGCAGCATGCCAGATACCTATCTGTGCTGGACAAATGGTTGCTGCCCCGGAACCCG 1276

Qy 1141 CCCAGTGAGGACCTCCCAGGGCCTGGTGGCAGTGTGGACATAGTGGCCATGGATGGAAGGC 1200
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1277 CCCAGTGAGGACCTCCCAGGGCCTGGTGGCAGTGTGGACATAGTGGCCATGGATGGAAGGC 1336

Qy 1201 TCAGAAGCATCCTCCTGCTTCGTCATCGCTTTGGCCTCGAAGCCTCGGAGCCCCAGAGAGGGAGCCTCT 1260
          |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1337 TCAGAAGCATCCTCCTGCTTCGTCATCGTCATCGCTTTGGCCTCGAAGCCTCGGAGCCCCAGAGGGGAGCCTCT 1396

Qy 1261 GCTGCCAGCTTTGAGTACATACTATCCTGACCATCCTGGCGTCTTCCCAGCTCTGCGTCCATGGACA 1320
          ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1397 GCTGCCAGCTTTGAGTACATACTATCCTGACCCCCAGCTCCAGCTCTGCGTCCATGGACA 1456

Qy 1321 CTGTGCCCTGAGCTGCCCTACCCCACCCCACCTAAAGTACCTGTACCTTGTGGTATCT 1380
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1457 CTGTGCCCTGAGCTGCCCTACCCCACCCCACCTAAAGTACCTGTACCTTGTGGTATCT 1516

Qy 1381 GACTCTGGCATCTCAACTGACTACAGCTCCAGGGGACTCCAGGGAGCCCAAGGGGGCTTA 1440
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1517 GACTCTGGCATCTCAACTGACTACAGCTCCAGGGGACTCCAGGGAGCCCAAGGGGGCTTA 1576

Qy 1441 TCCGATGGCCCCTACTTCCAACCCTTATGAGAACAGCCTTATCCCAGCCGCTGAGCCTCTG 1500
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1577 TCCGATGGCCCCTACTTCCAACCCTTATGAGAACAGCCTTATCCCAGCCGCTGAGCCTCTG 1636

Qy 1501 CCCCCCAGCTATGTGGCTTGCTCTTAG 1527
          |||||||||||||||||||||||||||
Db 1637 CCCCCCAGCTATGTGGCTTGCTCTTAG 1663

```
CC      - cythemia vera. See also AAQ05747 (murine EPO receptor clone). (Updated
CC        on 25-MAR-2003 to correct PA field.) (Updated on 25-MAR-2003 to correct
CC        PI field.)
XX
SQ      Sequence 1883 BP; 320 A; 607 C; 584 G; 372 T; 0 U; 0 Other;

Query Match     100.0%;  Score 1527;  DB 2;  Length 1883;
Best Local Similarity   100.0%;  Pred. No. 0;
Matches 1527;  Conservative   0;  Mismatches   0;  Indels   0;  Gaps   0;

Qy        1 ATGGACCACCTCGGGGCGGTCCTCTCGGCCCCAGTCGGTCCCTTTGTCTCCTTGCTCGCT   60
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      145 ATGGACCACCTCGGGGCGGTCCTCTCGGCCCCAGTCGGTCCCTTTGTCTCCTTGCTCGCT  204

Qy       61 GGGGCCCGCCTGGGCGCCCCCCGCTCAACCTCCCGGACCCCAAGTTCGAGAGCAAAGCGCC  120
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      205 GGGGCCCGCCTGGGCGCCCCCCGCTCAACCTCCCGGACCCCAAGTTCGAGAGCAAAGCGGCC  264

Qy      121 TTGCTGGCGGGCCCCGGGGCCCGAGAGCTTCTGTGCTTCACCGAGCGGTTGGAGGACTTG  180
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      265 TTGCTGGCGGGCCCCGGGGCCCGAGAGCTTCTGTGCTTCACCGAGCGGTTGGAGGACTTG  324

Qy      181 GTGTGTTTCTGGGAGGAAGCGGCGAGCGCTGGGGTGGGCCCGGGCCAACTACAGCTTCTCC  240
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      325 GTGTGTTTCTGGGAGGAAGCGGCGAGCGCTGGGGTGGGCCCGGGCCAACTACAGCTTCTCC  384

Qy      241 TACCAGCTCGAGGATGAGCCATGGAAGCTGTGTCGCCTACAGCCGACACGTCGAGCTTCGT  300
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      385 TACCAGCTCGAGGATGAGCCATGGAAGCTGTGTCGCCTACAGCCGACACGTCGAGCTTCGT  444

Qy      301 GGTGCGGTGGCGCTTCGGTGTTCGCTGCCTACAGCCGACCGACAGTCGAGCTTGTCGCCCCTA  360
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      445 GGTGCGGTGGCGCTTCGGTGTTCGCTGCCTACAGCCGACCGACAGTCGAGCTTGTCGCCCCTA  504

Qy      361 GAGTTGGCGGTCACAGCAGCCTTCCGGCCGTCCGGCGATATCACCGTGTCATCACATCAAT  420
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      505 GAGTTGGCGGTCACAGCAGCCTTCCGGCCGTCCGGCGATATCACCGTGTCATCACATCAAT  564

Qy      421 GAAGTAGTGCTCCTAGACGCCCCCGTGGGGCGCGGTGGTGCGGCGGTTGGCTGACGAGAGCGGC  480
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      565 GAAGTAGTGCTCCTAGACGCCCCCGTGGGGCGCGGTGGTGCGGCGGTTGGCTGACGAGAGCGGC  624

Qy      481 CACGTAGTGTTCGGCCGGCAACGTCTCGGCCGGAGCGTCACAGAGGTGGAGATCCGCTAC  540
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      625 CACGTAGTGTTCGGCCGGCAACGTCTCGGCCGGAGCGTCACAGAGGTGGAGATCCGCTAC  684

Qy      541 GAGGTGGACGTCTCGGCCGGCAACGGGCCAGGAGCGTACAGAGAGGTGGAGAGTCCTGGAG  600
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      685 GAGGTGGACGTCTCGGCCGGCAACGGGCCAGGAGCGTACAGAGAGGTGGAGAGTCCTGGAG  744

Qy      601 GGCCGCACCGAGTGTGTGCTGGACAACCTGCGCGGGCCGCCTACACCTTCGCGTC  660
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      745 GGCCGCACCGAGTGTGTGCTGGACAACCTGCGCGGGCCGCCTACACCTTCGCGTC  804

Qy      661 CGCCGCGCGTATGGCCGACCGAGCCGGAGCCTTCGGCCGGCCTTCGCGGAGCCGTG  720
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      805 CGCCGCGCGTATGGCCGACCGAGCCGGAGCCTTCGGCCGGCCTTCGCGGAGCCGTG  864

Qy      721 TCGCTGCTGACGGCCTAGCGACCTGGACCGCTCTTCATCCTGACGCTCCTCATCCTGTG  780
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      865 TCGCTGCTGACGGCCTAGCGACCTGGACCGCTCTTCATCCTGACGCTCCTCATCCTGTG  924

Qy      781 GTCATCCTGGGTCGCTGCTACCGTCTGCGCGTGCTCTCGCGCCTTCCCACCGCGGCTCGAAGCAG  840
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      925 GTCATCCTGGGTCGCTGCTACCGTCTGCGCGTGCTCTCGCGCCTTCCCACCGCGGCTCGAAGCAG  984

Qy      841 AAGATCTGGCCTGGCATCCCGAGCCCCAGAGAGCGAGTTTGAAGGCCTCTTCACCACCCAC  900
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      985 AAGATCTGGCCTGGCATCCCGAGCCCCAGAGAGCGAGTTTGAAGGCCTCTTCACCACCCAC  1044

Qy      901 AAGGGTAACTTCCAGTTGGCTGTGTACCAGAATGATGCTGCCTGGTGTGTGGTGAGCCCCTGC  960
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1045 AAGGGTAACTTCCAGTTGGCTGTGTACCAGAATGATGCTGCCTGGTGTGTGGTGAGCCCCTGC  1104
```

```
Qy      961 ACCCCCTTCACGGACGGAGGACCACCTGCTTCCCTGGAAGTCCTCTTCAGAGGCGCTCTCTGGGGG  1020
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1105 ACCCCCTTCACGGACGGAGGACCACCTGCTTCCCTGGAAGTCCTCTTCAGAGGCGCTCTCTGGGGG  1164

Qy     1021 ACGATGCAGGCAGTGAGGCCGGGACAGAGATGAGGGCCCCTGCTCTGGAGCGGTGGGC  1080
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1165 ACGATGCAGGCAGTGAGGCCGGGACAGAGATGAGGGCCCCTGCTCTGGAGCGGTGGGC  1224

Qy     1081 AGTGAGCATGCCCAGGATACCTATCTGGTGTGGACAAATGTTGTGCCCCGGAACCCG  1140
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1225 AGTGAGCATGCCCAGGATACCTATCTGGTGTGGACAAATGTTGTGCCCCGGAACCCG  1284

Qy     1141 CCCAGTGAGGACTCCGGTGGCAGTGTGGACATAGTGGCCATGGATGATGAAGGC  1200
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1285 CCCAGTGAGGACTCCGGTGGCAGTGTGGACATAGTGGCCATGGATGATGAAGGC  1344

Qy     1201 TCAGAAGCATCTCCTGCTCATCGTCTTGGCCTCGAAGCCCAGAAGGGAGCCTCT  1260
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1345 TCAGAAGCATCTCCTGCTCATCGTCTTGGCCTCGAAGCCCAGAAGGGAGCCTCT  1404

Qy     1261 GCTGCCAGCTTTGAGTACACTATCCTGGACCCCAGCTCTGCGTCCATGGACA  1320
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1405 GGTGCCAGCTTTGAGTACACTATCCTGGACCCCAGCTCTGCGTCCATGGACA  1464

Qy     1321 CTGTGCCCGAGCTGCCCCTACCCCACCCCACCTAAAGTACCTGTACCTTGTGGTATCT  1380
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1465 CTGTGCCCGAGCTGCCCCTACCCCACCCCACCTAAAGTACCTGTACCTTGTGGTATCT  1524

Qy     1381 GACTCTGCATCTCAACTGACTACAGTCAGGGGACTTCCCAGGGAGCCCAAGGGGCTTA  1440
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1525 GACTCTGCATCTCAACTGACTACAGTCAGGGGACTTCCCAGGGAGCCCAAGGGGCTTA  1584

Qy     1441 TCCGATGGCCCCTACTCCAACCTTATGAGAACAGCCTTATCCCAGCCCTGCTGAGCCCTCTG  1500
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1585 TCCGATGGCCCCTACTCCAACCTTATGAGAACAGCCTTATCCCAGCCCTGCTGAGCCCTCTG  1644

Qy     1501 CCCCCAGCTATGTGGCTTGCTCTTAG  1527
            ||||||||||||||||||||||||||
Db     1645 CCCCCAGCTATGTGGCTTGCTCTTAG  1671
```

```
RESULT 7
ABL51519
ID   ABL51519 standard; cDNA; 1527 BP.
XX
AC   ABL51519;
XX
DT   01-JUL-2002 (first entry)
XX
DE   Human erythropoietin receptor encoding cDNA SEQ ID NO:4.
XX
KW   Human; erythropoietin receptor; EPO receptor; fusion protein;
KW   erythropoietin receptor extracellular domain; signal transduction;
KW   receptor; gene; db.
XX
OS   Homo sapiens.
XX
FH   Key            Location/Qualifiers
FT   CDS            1..1527
FT                  /*tag= a
FT                  /product= "erythropoietin receptor"
XX
PN   US2002031806-A1
XX
PD   14-MAR-2002.
XX
PF   30-JAN-1998;  98US-00016159.
XX
PR   16-JUN-1997;  97US-00876227.
XX
PA   (LEEJ/) LEE J Y.
XX
```

```
PI   Lee JY;
XX
DR   WPI; 2002-361181/39.
DR   P-PSDB; ABB09173.
XX
PT   New purified human erythropoietin receptor polypeptide, useful for
PT   studying ligand binding to erythropoietin receptor, as well as for
PT   analyzing and understanding receptor structure and signal transduction.
XX
PS   Example 1; Page 9-10; 19pp; English.
XX
CC   The present invention describes a purified human erythropoietin (EPO)
CC   receptor polypeptide, which consists essentially of amino acids 25-250 of
CC   the full length human EPO receptor protein. The human EPO receptor
CC   polypeptide is capable of binding human erythropoietin. Also described is
CC   a purified fusion protein comprising essentially of: (a) a first
CC   polypeptide segment having an amino terminus and a carboxyl terminus, the
CC   segment having a thrombin proteolytic cleavage site at the carboxyl
CC   terminus; and (b) a second polypeptide segment consisting essentially of
CC   the purified human EPO receptor, where the second polypeptide segment is
CC   covalently coupled to the carboxyl terminus of the first polypeptide
CC   segment. The human EPO receptor polypeptide is useful for studying ligand
CC   binding to the EPO receptor and for quantitating the amounts of the EPO
CC   receptor. The polypeptide is also useful for analysing and understanding
CC   receptor structure and signal transduction. The present sequence encodes
CC   human EPO receptor from the present invention
XX
SQ   Sequence 1527 BP; 246 A; 514 C; 469 G; 298 T; 0 U; 0 Other;

Query Match             99.9%;  Score 1525.4;  DB 6;  Length 1527;
Best Local Similarity   99.9%;  Pred. No. 0;
Matches 1526;  Conservative    0;  Mismatches    0;  Indels    1;  Gaps    0;

Qy      1 ATGGACCACCTCGGGGCGTCCCTCTGGCCCCAGGTCGGCGTCCCTTGTCTCCTGCTCGCT 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 ATGGACCACCTCGGGGCGTCCCTCTGGCCCCAGGTCGGCGTCCCTTGTCTCCTGCTCGCT 60

Qy     61 GGGGCCGCCTGGGCGCCCCCCGCCCTAACCTCCCGGACCCCAAGTTGAGAGCAAAGCGGCC 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61 GGGGCCGCCTGGGCGCCCCCCGCCCTAACCTCCCGGACCCCAAGTTGAGAGCAAAGCGGCC 120

Qy    121 TTGCTGGCGGCCCGGGGGCCCGAAGAGCTTCTGTGCTTCACCGAGGCGGTTGGAGGACTTG 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121 TTGCTGGCGGCCCGGGGGCCCGAAGAGCTTCTGTGCTTCACCGAGGCGGTTGGAGGACTTG 180

Qy    181 GTGTGTTTCTGGGAGGAAGCGGCGAGCGGCTGGGGTGGCCCGGGCAACTACAGCTTCTCC 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    181 GTGTGTTTCTGGGAGGAAGCGGCGAGCGGCTGGGGTGGCCCGGGCAACTACAGCTTCTCC 240

Qy    241 TACCAGTCGAAGATGAGCCATGGAAGCTGTGTCGCCTACAGCGCACCTCCCACGGCTCGT 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    241 TACCAGTCGAAGATGAGCCATGGAAGCTGTGTCGCCTACAGCGCACCTCCCACGGCTCGT 300

Qy    301 GGTGCGGTGCGCGTTCGCGGCGCTTCGCCTACAGCCGCACAGCTGAGCTTCGTGCCCCTA 360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    301 GGTGCGGTGCGCGTTCGCGGCGCTTCGCCTACAGCCGCACAGCTGAGCTTCGTGCCCCTA 360

Qy    361 GAGTTGCGGTCACAGCAGCTCCGCGGCGTCCGGCGCTCCGGGCGCTGGTGCGGTGCGGGC 420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    361 GAGTTGCGGTCACAGCAGCTCCGCGGCGTCCGGCGCTCCGGGCGCTGGTGCGGTGCGGGC 420

Qy    421 GAAGTAGTGCTTCCTAGACGCCCCGTGGGCGTTGGCGTGACGAGAGCGGC 480
          ||||||||||||||||||||||||||||||||||||||||||||||||||
Db    421 GAAGTAGTGCTTCCTAGACGCCCCGTGGGCGTTGGCGTGACGAGAGCGGC 480

Qy    481 CACGTAGTGTTGCGCGTCTCGGCGCTGGCTCCGCCCGCCATGAGACGTCTCACATCCGCTAC 540
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    481 CACGTAGTGTTGCGCGTCTCGGCGCTGGCTCCGCCCGCCATGAGACGTCTCACATCCGCTAC 540

Qy    541 GAGGTCGGACGTCTCGGCGCCGGCAACGGCGCAGGGAGGGCTACAGAGAGGTGGAGATCCTGGAG 600
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    541 GAGGTCGGACGTCTCGGCGCCGGCAACGGCGCAGGGAGGGCTACAGAGAGGTGGAGATCCTGGAG 600
```

```
Qy    601 GGCCGCGACCGAGTGTGTGCTGAGCAACCTCGGGGGCCGAGCGGCTACACCTTCGCCGTC 660
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    601 GGCCGCGACCGAGTGTGTGCTGAGCAACCTCGGGGGCCGAGCGGCTACACCTTCGCCGTC 660

Qy    661 CGCGGCGGTATGGCTGAGCCGAGCTTCGGCGGCTTCGGCGACCCGTCTCCTCCATCCTCGTG 720
          | |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    661 CTCATCCTCGGTGCTGCGACCCGTCTCCTCCACCGCCGGGCTCTGGGAGCGCTCGAAGCAG 720

Qy    721 TCGCTGCTGACGCTAGCGACCTGGACCCCTGGACCCCCTCATCCTCGACGCTCTCATCCTCGTG 780
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    721 TCGCTGCTGACGCTAGCGACCTGGACCCCTGGACCCCCTCATCCTCGACGCTCTCATCCTCGTG 780

Qy    781 GTCATCCTGGTGCTGCTGAGCCGTCGCGTCGACCGGTCGCGCTGCTCCCACCGCCGGGCTCTGAAGCAG 840
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    781 GTCATCCTCGGTGCTGCTGAGCCGTCGCGTCGACCGGTCGCGCTGCTCCCACCGCCGGGCTCTGAAGCAG 840

Qy    841 AAGATCTGGCTGGCATCCCGAGCCCAGAGAGCGAGTTGAAGGCCTTCACCACCCAC 900
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    841 AAGATCTGGCTGGCATCCCGAGCCCAGAGAGCGAGTTGAAGGCCTTCACCACCCAC 900

Qy    901 AAGGGTAACTTCCAGCTGTGGCTGTACCAGGATGATGCTGCTGTGTGGGAGCCCCTGC 960
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    901 AAGGGTAACTTCCAGCTGTGGCTGTACCAGGATGATGCTGCTGTGTGGGAGCCCCTGC 960

Qy    961 ACCCCCTCACGGAGACCCACCTGCTTCCCTGGAGTCCTCTCAGAGACGTCCTCTCAGAGCGCTGTGGGG 1020
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    961 ACCCCCTCACGGAGACCCACCTGCTTCCCTGGAGTCCTCTCAGAGACGTCCTCTCAGAGCGCTGTGGGG 1020

Qy   1021 ACGATGCAGGCAGTGAGCAGTGGACAGATGATGAGGGCCCCTGCTGGAGCCAGTGGGC 1080
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1021 ACGATGCAGGCAGTGAGCAGTGGACAGATGATGAGGGCCCCTGCTGGAGCCAGTGGGC 1080

Qy   1081 AGTGAGCATACCTATCTGTGCTGGACAAATGGTTGCTGCCCCGGAACCCG 1140
          ||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1081 AGTGAGCATACCTATCTGTGCTGGACAAATGGTTGCTGCCCCGGAACCCG 1140

Qy   1141 CCCAGTGAGGAGCTCCCAGGGCCTGGTGGACATAGTGGGCACATAGTGGACATAGTGGAAGGC 1200
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1141 CCCAGTGAGGAGCTCCCAGGGCCTGGTGGACATAGTGGGCACATAGTGGACATAGTGAAGGC 1200

Qy   1201 TCAGAAGCATCCTCCTGCTCATCGTCTTTGGCCTCGAAGCCCAGCCCAGAGGAGGAGCCTCT 1260
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1201 TCAGAAGCATCCTCCTGCTCATCGTCTTTGGCCTCGAAGCCCAGCCCAGAGGAGGAGCCTCT 1260

Qy   1261 GCTGCCAGCTTTGAGTACACTATCCTGGACCCCCAGTCTTGCGTCTTCCATGGACA 1320
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1261 GCTGCCAGCTTTGAGTACACTATCCTGGACCCCCAGTCTTGCGTCTTCCATGGACA 1320

Qy   1321 CTGTGCCTGAGCTGCCCCCTACCCACCTAAAGTACCTGTACCTTGTGGTATCT 1380
          ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1321 CTGTGCCTGAGCTGCCCCCTACCCACCTAAAGTACCTGTACCTTGTGGTATCT 1380

Qy   1381 GACTCTGGCATCTCAACTGACTACAGCTCAGGGGAGGACTCCCAAGGGGGCTTA 1440
          ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1381 GACTCTGGCATCTCAACTGACTACAGCTCAGGGGAGGACTCCCAAGGGGGCTTA 1440

Qy   1441 TCCGATGGCCCCTACTTCAAACCTTATGAGAACAGCCTTATCCAGCCGCTCTCTG 1500
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1441 TCCGATGGCCCCTACTTCAAACCTTATGAGAACAGCCTTATCCAGCCGCTCTCTG 1500

Qy   1501 CCCCCCAGCTATGTGGCTTGCTTCTTAG 1527
          |||||||||||||||||||||||||||
Db   1501 CCCCCCAGCTATGTGGCTTGCTTCTTAG 1527
```

```
RESULT 8
AAQ81892
ID   AAQ81892 standard; cDNA; 1818 BP.
XX
AC   AAQ81892;
XX
DT   25-MAR-2003 (revised)
DT   11-AUG-1995 (first entry)
```

```
XX
DE   Human erythropoietin receptor cDNA.
XX
KW   Erythropoietin receptor; cDNA; anemia therapy; diagnostic; ds.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   CDS             106..1632
FT                   /product= "Erythropoietin receptor"
FT   exon            106..220
FT                   /*tag= a
FT                   /*tag= b
FT   sig_peptide     106..177
FT                   /*tag= k
FT   mat_peptide     178..1629
FT                   /*tag= l
FT   exon            221..356
FT                   /*tag= c
FT   conflict.       341..341
FT                   /*tag= m
FT                   /note "base difference: A in fetal liver clone"
FT   exon            357..532
FT                   /*tag= d
FT   exon            533..690
FT                   /*tag= e
FT   exon            691..850
FT                   /*tag= f
FT   exon            851..932
FT                   /*tag= g
FT   exon            933..1021
FT                   /*tag= h
FT   exon            1022..1632
FT                   /*tag= i
XX
PN   US5378808-A.
XX
PD   03-JAN-1995.
XX
PF   10-JUN-1993;  93US-00075069.
XX
PR   03-FEB-1989;  89US-00306503.
PR   25-MAR-1991;  91US-00678877.
XX
PA   (GEMY ) GENETICS INST INC.
XX
PI   Jones SS,  D'andrea A,  Wong GG;
XX
DR   WPI; 1995-051310/07.
DR   P-PSDB; AAR69503.
XX
PT   recombinant erythropoietin receptor polypeptide(s) - used for detection,
PT   purificn. and therapy and for prodn. of antibodies for detection and
PT   therapy.
XX
PS   Claim 2; Fig 9; 24pp; English.
XX
CC   The sequence encodes a human erythropoietin receptor cDNA (clone 18).
CC   Exon boundaries and a base difference with the PCR clone have been
CC   determined. The gene product may be used for erythropoietin production or
CC   detection, and in production of antibodies for anemia therapy. The human
CC   gene has been isolated from a DNA library using a mouse DNA probe.
CC   (Updated on 25-MAR-2003 to correct PF field.)
XX
SQ   Sequence 1818 BP; 308 A; 585 C; 565 G; 360 T; 0 U; 0 Other;

Query Match      99.9%;  Score 1525.4;  DB 2;  Length 1818;
Best Local Similarity   99.9%;
Matches 1526;  Conservative  0;  Mismatches  0;  Indels  1;  Gaps  0;
```

```
Qy      1 ATGGACCACCTCGGGGCGTCCCTCTGGCCCCAGGTCGGCCTCCCTTTGTCTCCGCTCGCT    60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    106 ATGGACCACCTCGGGGCGTCCCTCTGGCCCCAGGTCGGCCTCCCTTTGTCTCCGCTCGCT   165

Qy     61 GGGGCCCCTGGGCGGCCCCCGGCCTAACCTCCCGGACCCCCAAGTTCGAGAGCAAAGGCGGCC   120
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    166 GGGGCCCCTGGGCGGCCCCCGGCCTAACCTCCCGGACCCCCAAGTTCGAGAGCAAAGGCGGCC   225

Qy    121 TTGCTGGCGGCCCCGGGGGCCCGGAAGAGAGCTTCTGCTTCGTGCTTCACCGAGCGGTTGGAGGACTTG   180
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    226 TTGCTGGCGGCCCCGGGGGCCCGGAAGAGAGCTTCTGCTTCGTGCTTCACCGAGCGGTTGGAGGACTTG   285

Qy    181 GTGTGTTTCTGGGAGGAGCGGCGAGCGCTGGGCGTGGCGGCCGCAACTACAGCTTCTCC   240
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    286 GTGTGTTTCTGGGAGGAGCGGCGAGCGCTGGGCGTGGCGGCCGCAACTACAGCTTCTCC   345

Qy    241 TACCAGCTGAGGATGAGCAATGGAAGCTGTGTCGCCTGCACCAGGTCCCACGGCTCGT   300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    346 TACCAGCTGAGGATGAGCCATGGAAGCTGTGTCGCCTGCACCAGGTCCCACGGCTCGT   405

Qy    301 GGTGCGGTGCGCTTCTGTGTTCGCTGCCTCACAGCCGACAGCGACACACGTCGAGCTTCGTGCCCCTA   360
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    406 GGTGCGGTGCGCTTCTGTGTTCGCTGCCTCACAGCCGACAGCCGACACACGTCGAGCTTCGTGCCCCTA   465

Qy    361 GAGTTGCGGTCACACGAGCCTCCGGCGCTCCGGCGATATCACCGTGTCATCCACATCAAT   420
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    466 GAGTTGCGGTCACACGAGCCTCCGGCGCTCCGGCGATATCACCGTGTCATCCACATCAAT   525

Qy    421 GAAGTAGTGCTCCTAGACGCCCCCGTGGGCGTTGGTGGCGGCCGGT GGCTGCGTGACGAGAGCGGC   480
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    526 GAAGTAGTGCTCCTAGACGCCCCCGTGGGCGTTGGTGGCGGCCGGT GGCTGCGTGACGAGAGCGGC   585

Qy    481 CACGTAGTGTTGGCTGCCTGCTCCGAGACGTCTCACATCCGCTAC   540
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    586 CACGTAGTGTTGGCTGCCTGCTCCGAGACGTCTCACATCCGCTAC   645

Qy    541 GAGGTGACGTCTCGGCGCCGGCAACGGCGCAGGAGGCGCAGAGGGTGGAGATCCTGAG   600
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    646 GAGGTGACGTCTCGGCGCCGGCAACGGCGCAGGAGGCGCAGAGGGTGGAGATCCTGAG   705

Qy    601 GGCCGCACCGAGTGTGCTGCTGAGCAACCTGGCCGACGCGCTACACCTTCGCCGTC   660
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    706 GGCCGCACCGAGTGTGCTGCTGAGCAACCTGGCCGACGCGCTACACCTTCGCCGTC   765

Qy    661 CGGCGGCGTATGGCTGAGCCGAGCTTCGGAGCCTCGTGTGGAGCCTGTG   720
          ||||||||||||||||||||||||||||||||||||||||||||||||||
Db    766 CGGCGGCGTATGGCTGAGCCGAGCTTCGGAGCCTCGTGTGGAGCCTGTG   825

Qy    721 TCGCTGCTGAGGCCTAGCGACCTGGACCCCCTCATCCTGACGCTCTCCCTCATCCTCGTG   780
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    826 TCGCTGCTGAGGCCTAGCGACCTGGACCCCCTCATCCTGACGCTCTCCCTCATCCTCGTG   885

Qy    781 GTCATCCTGGTGTCGTGCTGCTGAGCGTGCTGACCGCGTCCCCACCGCCCACGGCTCTGAAGCAG   840
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    886 GTCATCCTGGTGTCGTGCTGCTGAGCGTGCTGACCGCGTCCCCACCGCCCACGGCTCTGAAGCAG   945

Qy    841 AAGATCTGGCCATCCCGAGCCCAGAGAGCGAGTTGAAGGCCTTCACCCACCACCCAC   900
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    946 AAGATCTGGCCATCCCGAGCCCAGAGAGCGAGTTGAAGGCCTTCACCCACCACCCAC   1005

Qy    901 AAGGGTAACTTCCAGCTGTGGCTGTACCAGATGATGGCTGCTGTGCTGAGCCCCTGC   960
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1006 AAGGGTAACTTCCAGCTGTGGCTGTACCAGAATGATGGCTGCTGTGCTGAGCCCCTGC   1065

Qy    961 ACCCCCTTCACCGGAGGACCCACCTGCTTCCCTCAGAGCCTGTCTGGGGG   1020
          |||||||||||||||||||||||||||||||||||||||||||||||||
Db   1066 ACCCCCTTCACCGGAGGACCCACCTGCTTCCCTCAGAGCCTGTCTGGGGG   1125

Qy   1021 ACGATGCAGGCAGTGGAGCCCGGGGACAGATGAGTGAGGGCCCCCTGCTGTGGGAGCCAGTGGGC   1080
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1126 ACGATGCAGGCAGTGGAGCCCGGGGACAGATGAGTGAGGGCCCCCTGCTGTGGGAGCCAGTGGGC   1185

Qy   1081 AGTGAGCATGCCCAGGATACCTATCTGGTGCTGGACAAATGGTTGCTCCCCCGGAACCCG   1140
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1186 AGTGAGCATGCCCAGGATACCTATCTGGTGCTGGACAAATGGTTGCTCCCCCGGAACCCG   1245
```

```
Qy   1141  CCCAGTGAGGACCTCCAGGGCCTGGTCGGCAGTGTGGACATAGTGGCCATGATGAAGGC  1200
                  ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1246  CCCAGTGAGGACCTCCCAGGGCCTGGTCGGCAGTGTGGACATAGTGGCCATGATGAAGGC  1305

Qy   1201  TCAGAAGCATCCTCCTGCTCATCTGCTTGGCCTCGAAGCCCAGCCCAGAGGGAGCCTCT  1260
                  ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1306  TCAGAAGCATCCTCCTGCTCATCTGCTTGGCCTCGAAGCCCAGCCCAGAGGGAGCCTCT  1365

Qy   1261  GCTGCCAGCTTTGAGTACATTACATATCCTGACACCCAGCTCCGAGTCTTGCGTCCATGGACA  1320
                  ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1366  GCTGCCAGCTTTGAGTACACTACTATCCTGACCACCCCAGCTCCGAGTCTTGCGTCCATGGACA  1425

Qy   1321  CTGTGCCCTGAGCTGCCCCCTACCCCACCCCACCTAAAGTACCTGTACCTTGTGGTATCT  1380
                  ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1426  CTGTGCCCTGAGCTGCCCCCCTACCCCCACCCCCCACCTAAAGTACCTGTACCTTGTGGTATCT  1485

Qy   1381  GACTCTGCATCTCAACTGACTACAGCTCAGGGGACTCCCAGGGAGGCCAAGGGGCTTA  1440
                  ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1486  GACTCTGCATCTCAACTGACTACAGCTCAGGGGACTCCCAGGGAGGCCAAGGGGCTTA  1545

Qy   1441  TCCGATGGCCCCTACTCCAACCCTTATGAGAACAGCCTTATCCCAGCCGCTGAGCCTCTG  1500
                  ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1546  TCCGATGGCCCCCTACTCCAACCCTTATGAGAACAGCCTTATCCCAGCCGCTGAGCCTCTG  1605

Qy   1501  CCCCCAGCTATGTGGCTTGCTCTTAG  1527
                  ||||||||||||||||||||
Db   1606  CCCCCCAGCTATGTGGCTTGCTCTTAG  1632
```

RESULT 9
ADE28676
ID   ADE28676 standard; cDNA; 1585 BP.
XX
AC   ADE28676;
XX
DT   29-JAN-2004 (first entry)
XX
DE   Human NOV15d cDNA - SEQ ID 53.
XX
KW   NOVX; antidiabetic; anorectic; cardiant; hypotensive;
KW   antiarteriosclerotic; virucide; antibacterial; fungicide; protozoacide;
KW   nootropic; neuroprotective; antiparkinsonian; anticonvulsant;
KW   osteopathic; antiarthritic; antiinflammatory; dermatological;
KW   antiasthmatic; antilipaemic; metabolic; diabetes; obesity; infectious;
KW   anorexia; cancer; cardiovascular; hypertension; atherosclerosis;
KW   neurodegenerative; Alzheimer's disease; Parkinson's disease; epilepsy; immune;
KW   osteoarthritis; haemopoietic; inflammatory skin; asthma; dyslipidaemia;
KW   neurogenesis; cell differentiation; proliferation; haemopoiesis;
KW   wound healing; angiogenesis; gene therapy; NOV.
KW   tissue typing; human; ss; gene; NOV.
XX
OS   Homo sapiens.
XX
PN   WO200304030-A2.
XX
PD   15-MAY-2003.
XX
PF   05-NOV-2002; 2002WO-US035536.
XX
PR   05-NOV-2001; 2001US-0338626P.
PR   05-DEC-2001; 2001US-0336600P.
PR   07-DEC-2001; 2001US-0338285P.
PR   12-DEC-2001; 2001US-0341146P.
PR   17-DEC-2001; 2001US-0341477P.
PR   17-DEC-2001; 2001US-0341540P.
PR   20-DEC-2001; 2001US-0342592P.
PR   27-DEC-2001; 2001US-0344297P.
PR   31-DEC-2001; 2001US-0344903P.
PR   17-APR-2002; 2002US-0373288P.
PR   15-MAY-2002; 2002US-0380981P.
PR   17-MAY-2002; 2002US-0381495P.
PR   28-MAY-2002; 2002US-0383534P.
PR   28-MAY-2002; 2002US-0383744P.
PR   29-MAY-2002; 2002US-0383829P.
PR   29-MAY-2002; 2002US-0384024P.
PR   07-AUG-2002; 2002US-0401788P.
PR   26-AUG-2002; 2002US-0406353P.
PR   31-OCT-2002; 2002US-00287971.
XX
PA   (CURA-) CURAGEN CORP.
XX
PI   Alsobrook JP, Alvarez E, Anderson DW, Baron M, Boldog FL;
PI   Burgess CE, Casman SJ, Chapoval A, Dhanabal M, Edinger SR, Eisen A;
PI   Ellerman K, Ettenberg S, Gangolli EA, Gerlach VL, Gorman L;
PI   Grosse WM, Guo X, Hackett C, Ji W, Kekuda R, Khramtsov NV;
PI   Lepley DM, Li L, Macdougall JR, Malyankar UM, Mazur A, Mcqueeney K;
PI   Mezes PS, Miller CE, Millet I, Mishra VS, Padigaru M, Patturajan M;
PI   Pena CEA, Peyman JA, Rastelli L, Rieger DK, Shenoy SG, Shimkets RA;
PI   Smithson G, Starling G, Spytek KA, Stone DJ, Tchernev VT, Twomlow N;
PI   Vernet CAM, Zerhusen BD, Zhong M;
XX
DR   WPI; 2003-441555/41.
DR   P-PSDB; ADE28677.
XX
PT   New isolated NOVX polypeptides and polynucleotides, useful for
PT   preventing, diagnosing or treating NOVX-associated disorders, e.g.
PT   osteoarthritis, obesity, atherosclerosis, cancer, Parkinson's disease,
PT   asthma, or infections.
XX
PS   Claim 20; SEQ ID NO 53; 447pp; English.
XX
CC   The invention relates to a novel isolated NOVX polypeptide. The
CC   polypeptide of the invention demonstrates, antidiabetic, anorectic,
CC   cardiant, hypotensive, antiarteriosclerotic, virucide, antibacterial,
CC   fungicide, protozoacide, nootropic, neuroprotective, antiparkinsonian,
CC   anticonvulsant, osteopathic, antiarthritic, antiinflammatory,
CC   dermatological, antiasthmatic and antilipaemic activities. The
CC   polypeptides, nucleic acid molecules and antibodies may be useful for
CC   treating or diagnosing diseases including metabolic disorders such as
CC   diabetes and obesity, infectious diseases, anorexia, cancer,
CC   cardiovascular diseases including hypertension and atherosclerosis,
CC   neurodegenerative disorders such as Alzheimer's disease, Parkinson's
CC   disease and epilepsy, immune disorders e.g. osteoarthritis, haemopoietic
CC   disorders, inflammatory skin disorders, asthma and dyslipidaemia.
CC   Furthermore, the nucleic acids and polypeptides may also be used to
CC   identify molecules that modulate or inhibit neurogenesis, cell
CC   differentiation and proliferation, haemopoiesis, wound healing and
CC   angiogenesis, as well as in gene therapy. Finally, the nucleic acids may
CC   be used as hybridisation probes, in chromosome mapping, tissue typing,
CC   preventive medicine and pharmacogenomics. The current sequence is that of
CC   the human NOV cDNA of the invention.
XX
SQ   Sequence 1585 BP; 263 A; 531 C; 486 G; 305 T; 0 U; 0 Other;

Query Match   99.6%; Score 1520.6; DB 10; Length 1585;
Best Local Similarity   99.7%;
Matches 1523; Conservative   0; Mismatches   4; Indels   0; Gaps   0;

```
Qy     1  ATGGACCACCTCGGGCGTCCCTCGCCCCAGGTCGGCTCCCCTTGTCTCCTTGTCTGCTCGCT   60
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    12  ATGGACCACCTCGGGCGTCCCTCGCCCCAGGTCGGCTCCCCTTGTCTCCTTGTCTGCTCGCT   71

Qy    61  GGGGCCGCTCGGCGCCTGGCGCCCCGGGGGCCCGAAGTTCGAGAGCAAAGCCGCC  120
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    72  GGGGCCGCTCGGCGCCTGGCGCCCCGGGGGCCCGAAGTTCGAGAGCAAAGCCGCC  131

Qy   121  TTGCTGGCGGCCCCGGGGCCCGAAGAGCTTCGTGTGCTTCACCGAGCGGTTGGAGGACTTG  180
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   132  TTGCTGGCGGCCCCGGGGCCCGAAGAGCTTCGTGTGCTTCACCGAGCGGTTGGAGGACTTG  191

Qy   181  GTGTGTTTCTGGAGGAAGCGGCGAGCGCGTGGGGTGGGCCCCGGGCAACTACAGCTTCTCC  240
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   192  GTGTGTTTCTGGAGGAAGCGGCGAGCGCGTGGGGTGGGCCCCGGGCAACTACAGCTTCTCC  251

Qy   241  TACCAGCTCGAGGATGAGCCATGGAAGCTGTGTCGCGCTGCGAAGCTGTGTCGCCCACCGGTCGTCGT  300
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db   1381 GACTCTGCGCATCTCAACTGACTACAGCTTCAGGGGACTCCCAGGGAGCCCAAGGGGGCTTA 1440
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1392 GACTCTGCGCATCTCAACTGACTACAGCTTCAGGGGACTCCCAGGGAGCCCAAGGGGGCTTA 1451

Qy   1441 TCCGATGGCCCCTACTCCAACCCTTATGAGAACAGCCCTTATCCGCTGAGCCGCTCTG 1500
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1452 TCCGATGGCCCCTACTCCAACCCTTATGAGAACAGCCCTTATCCGCTGAGCCGCTCTG 1511

Qy   1501 CCCCCCAGCTATGTGCTTGCTTCTTAG 1527
          |||||||||||||||||||||||||||
Db   1512 CCCCCCAGCTATGTGGCTTGCTCTTAG 1538


RESULT 10
ADM93421
ID   ADM93421 standard; cDNA; 1585 BP.
XX
AC   ADM93421;
XX
DT   01-JUL-2004  (first entry)
XX
DE   Human NOVX polynucleotide #27.
XX
KW   Human; NOVX; gene; ss; congenital heart defect; cardiomyopathy;
KW   atherosclerosis; hypertension; pulmonary stenosis; scleroderma;
KW   adenocarcinoma; haemophilia; graft-versus-host disease; cancer;
KW   neurodegenerative disorder; Alzheimer's disease; Parkinson's disease;
KW   multiple sclerosis; diabetes; obesity; bronchial asthma;
KW   acquired immunodeficiency syndrome; AIDS; Crohn's disease;
KW   infectious disease; anorexia; immune disorder.
XX
OS   Homo sapiens.
XX
PN   US2004067882-A1.
XX
PD   08-APR-2004.
XX
PF   05-NOV-2002;  2002US-00287971.
XX
PR   22-OCT-2001;  2001US-00035568.
PR   05-NOV-2001;  2001US-0338626P.
PR   06-NOV-2001;  2001US-0333072P.
PR   09-NOV-2001;  2001US-0345398P.
PR   09-NOV-2001;  2001US-0348283P.
PR   15-NOV-2001;  2001US-0335610P.
PR   21-NOV-2001;  2001US-0332152P.
PR   28-NOV-2001;  2001US-0333912P.
PR   29-NOV-2001;  2001US-00997425.
PR   29-NOV-2001;  2001US-0334300P.
PR   04-DEC-2001;  2001US-0336576P.
PR   05-FEB-2002;  2002US-0354807P.
PR   15-MAY-2002;  2002US-0380968P.
PR   16-MAY-2002;  2002US-0381043P.
PR   02-JUL-2002;  2002US-0393148P.
PR   02-JUL-2002;  2002US-0393262P.
PR   06-AUG-2002;  2002US-0401479P.
PR   06-AUG-2002;  2002US-0401626P.
PR   07-AUG-2002;  2002US-0401593P.
PR   07-AUG-2002;  2002US-0401695P.
PR   26-AUG-2002;  2002US-0406181P.
XX
PA   (ALSO/)  ALSOBROOK J P.
PA   (ALVA/)  ALVAREZ E.
PA   (ANDE/)  ANDERSON D W.
PA   (BARO/)  BARON M.
PA   (BOLD/)  BOLDOG F L.
PA   (BURG/)  BURGESS C E.
PA   (CASM/)  CASMAN S J.
PA   (CHAP/)  CHAPOVAL A.
PA   (DHAN/)  DHANABAL M.
PA   (EDIN/)  EDINGER S R.
PA   (EISE/)  EISEN A.
```

```
Db    252 TACCAGCTCGAGGATGAGCCATGGAAGCTGTGTCGCCTTGCACCGGCTCGT 311

Qy    301 GGTGCGGGTGCGCTTCTGGTTCGTGCTGCTACAGCCGACACGTCGGCCCCTA 360
          | ||||||| |||||||||||| ||||||||||||||| |||||||||||
Db    312 GGTGCGGGTGCGCGTTCTGGTTCGTGCTGCTACAGCCGACACGTCGTGCCCCTA 371

Qy    361 GAGTTGGCGGTCACAGACAGCCTTCGGGGCTCCGGCGGCCTCCGGCGGCGCTCCGCGATCACCGTCATCACAATCAAT 420
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    372 GAGTTGGCGGTCACAGACAGCCTTCGGGGCTCCGGCGGCCTCCGCGATCACCGTCATCACAATCAAT 431

Qy    421 GAAGTAGTGCTCCTAGACGCGCCCCCGTGGGCGTTGGTGGCGTGACGAGAGCGGGC 480
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    432 GAAGTAGTGCTCCTAGACGCGCCCCCGTGGGCGTTGGCTGGCGTGACGAGAGCGGGC 491

Qy    481 CACGTAGTGTTGCGGCTGCTGCTCCCGCCGCTGAGACGTGCTCACATCCGTAC 540
          |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    492 CACGTAGTGTTGCGGCTGCTGCTCCCGCCGCTGAGACACCCATGACGTCTCACATCCGTAC 551

Qy    541 GAGGTGGACGTCGCGGCCGCAACGGGCAGGAGCGTACAGAGGGTGGAGATCTGGAG 600
          |  |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    552 GCGGTGGACGTCGCGGCCGCAACGGGCAGGAGCGTACAGAGGGTGAAGATCCTGGAG 611

Qy    601 GGCCGCACCGAGTGTGTGCTGAGCGACAACCTGCCGGGGCCGGACGCGCCTTCGCCGTC 660
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    612 GGCCGCACCGAGTGTGTGCTGAGCGACAACCTGCCGGGGCCGGACGCGCCTACACCTTCGCCGTC 671

Qy    661 CGCGCGGGTATGCGTGCTGAGCCGGCCGAGCTTCGCGGCGCTTGGGAGCGCCGTGTCGGAGCCTGTG 720
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    672 CGCGCGGGTATGCGTGCTGAGCCGGCCGAGCTTCGCGGCGCTTGGGAGCGCCGTGTCGGAGCCTGTG 731

Qy    721 TCGCTGCTGACGCGCTAGCGACCTGACCTGACCACCCCCTCATCCTGACCCCTCATCCTCCCTGCTG 780
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    732 TCGCTGCTGACGCGCTAGCGACCTGACCTGACCACCCCCTCATCCTGACCCCTCATCCTCGTG 791

Qy    781 GTCATCCTGGTGCTGCTGACCGTGCTGCTCTCCCCGGGCTCTCCGCCGCTCTGAGCAG 840
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    792 GTCATCCTGGTGCTGCTGACCGTGCTGCTCTCCCCGGGCTCTCCGCCGCTCTGAAGCAG 851

Qy    841 AAGATCTGGCCCTGGCCATCCCAGAGCGCGAGTTTGAAGGCCTTCACCACCCAC 900
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    852 AAGATCTGGCCCTGGCCATCCCAGAGCGCGAGTTTGAAGGCCTTCACCACCCAC 911

Qy    901 AAGGGTAACTTCAGCTGTGGCTGTACCAGAATGATGGCTGCTGTGGTGGAGCCCCTGC 960
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    912 AAGGGTAACTTCAGCTGTGGCTGTACCAGAATGATGGCTGCTGTGGTGGAGCCCCTGC 971

Qy    961 ACCCCTTCACGGAGGACACCCACCTGCTTCCCTGGAGCTCCTCTCCAGAGCGCTGCTGGGGG 1020
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    972 ACCCCCTTCACGGAGGACACCCACCTGCTTCCCTGGAGCTCCTCTCCAGAGCGCTGCTGGGGG 1031

Qy   1021 ACGATGCAGTGGACAGCCGGGGACAGATGAGGGCCCCGTGCTGGAGCCAGTGGGGC 1080
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1032 ACGATGCAGGCAGTGGACAGATGAGGGCCCCGTGCTGGAGCCAGTGGGGC 1091

Qy   1081 AGTGAGCATGCCAGGATACCTATCTGTGCTGGACAAATGGTTGCTGCCCCGGAACCCG 1140
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1092 AGTGAGCATGCCAGGATACCTATCTGTGCTGGACAAATGGTTGCTGCCCCGGAACCCG 1151

Qy   1141 CCCAGTGAGGAGCTGGTGGCAGTGTGGCAGTAGTGGCCATGGATGAAGGC 1200
          ||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1152 CCCAGTGAGGAGCTGGTGGCAGTGTGGCAGTAGTGGCCATGGATGAAGGC 1211

Qy   1201 TCAGAAGCATCCTCCTGCTTTGGCCTCGAAGCCCCAGAGAGGAGGAGCCTCT 1260
          |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1212 TCAGAAGCATCCTCCTGCTTTGGCCTCTTTGGCCTCGAAGCCCCAGAGAGGGAGCCTCT 1271

Qy   1261 GCTGCCCAGCTTTGAGTACACTATCCTGGACCCCCAGCTTCCCAGCTCCATGGACA 1320
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1272 GCTGCCCAGCTTTGAGTACACTATCCTGGACCCCCAGCTCTTTGCCTCCATGGACA 1331

Qy   1321 CTGTGCCCGAGCTGCCCCCTACCCCACCCCACCCCTAAAGTACCTGTACCTTGTGTGGTTATCT 1380
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1332 CTGTGCCCGAGCTGCCCCCTACCCCACCCCACCCCTAAAGTACCTGTACCTTGTGTGGTTATCT 1391
```

```
PA   (ELLE/)   ELLERMAN K.
PA   (ETTE/)   ETTENBERG S.
PA   (GANG/)   GANGOLLI E A.
PA   (GERL/)   GERLACH V.
PA   (GORM/)   GORMAN L.
PA   (GROS/)   GROSSE W M.
PA   (GUOX/)   GUO X.
PA   (HACK/)   HACKETT C.
PA   (JIWW/)   JI W.
PA   (KEKU/)   KEKUDA R.
PA   (KHRA/)   KHRAMTSOV N V.
PA   (LEPL/)   LEPLEY D M.
PA   (LILL/)   LI L.
PA   (MACD/)   MACDOUGALL J R.
PA   (MALY/)   MALYANKAR U M.
PA   (MAZU/)   MAZUR A.
PA   (MCQU/)   MCQUEENEY K.
PA   (MEZE/)   MEZES P S.
PA   (MILL/)   MILLER C E.
PA   (MILL/)   MILLET I.
PA   (MISH/)   MISHRA V.
PA   (PADI/)   PADIGARU M.
PA   (PATT/)   PATTURAJAN M.
PA   (PENA/)   PENA C E A.
PA   (PEYM/)   PEYMAN J A.
PA   (RAST/)   RASTELLI L.
PA   (RIEG/)   RIEGER D K.
PA   (ROTH/)   ROTHENBERG M E.
PA   (SHEN/)   SHENOY S G.
PA   (SHIM/)   SHIMKETS R A.
PA   (SMIT/)   SMITHSON G.
PA   (SPAD/)   SPADERNA S K.
PA   (STAR/)   STARLING G.
PA   (SPYT/)   SPYTEK K A.
PA   (STON/)   STONE D J.
PA   (TCHE/)   TCHERNEV V T.
PA   (TWOM/)   TWOMLOW N.
PA   (VERN/)   VERNET C A M.
PA   (ZERH/)   ZERHUSEN B D.
PA   (VOSS/)   VOSS E Z.
PA   (ZHON/)   ZHONG M.
XX
PI   Alsobrook JP,  Alvarez E,  Anderson DW,   Baron M,    Boldog FL;
PI   Burgess CE,   Casman SJ,  Chapoval A,  Dhanabal M,  Edinger SR,   Eisen A;
PI   Ellerman K,   Ettenberg S,  Gangolli EA,  Gerlach v,   Gorman L;
PI   Grosse WM,   Guo X,   Hackett C,  Ji W, Kekuda R,   Khramtsov NV;
PI   Lepley DM,   Li L, Macdougall JR, Malyankar UM,  Mazur A,  Mcqueeney K;
PI   Mezes PS,   Miller CE,  Millet I,  Mishra V,  Padigaru M,  Patturajan M;
PI   Pena CEA,   Peyman JA,  Rastelli L,  Rieger DK,  Rothenberg ME;
PI   Shenoy SG,   Shimkets RA,  Smithson G,  Spaderna SK,  Starling G;
PI   Spytek KA,   Stone DJ,  Tchernev VT,  Twomlow N,  Vernet CAM;
PI   Zerhusen BD,  Voss EZ,   Zhong M;
XX
DR   WPI; 2004-355303/33.
DR   P-PSDB; ADM93422.
XX
PT   Novel isolated NOVX polypeptide useful treating or preventing disorders
PT   or syndromes such as Alzheimer's disease, Parkinson's disease, multiple
PT   sclerosis, diabetes, obesity, cancer, bronchial asthma, Crohn's disease.
XX
PS   Claim 20; SEQ ID NO 53; 330pp; English.
XX
CC   The invention relates to human NOVX polypeptides and the polynucleotides
CC   encoding them. The NOVX polypeptides and polynucleotides are useful for
CC   determining the presence of or predisposition to a disease associated
CC   with altered levels of the sequences in a mammalian subject, and for
CC   treating or preventing a pathology associated with NOVX. The
CC   polypeptides, polynucleotides and antibodies that bind immunospecifically
CC   to the polypeptides are useful for treating or preventing disorders or
CC   syndromes such as congenital heart defects, cardiomyopathy,
CC   atherosclerosis, hypertension, pulmonary stenosis, scleroderma,
CC   adenocarcinoma, haemophilia, graft-versus-host disease, cancer,
CC   neurodegenerative disorders, Alzheimer's disease, Parkinson's disease,
```

```
CC   multiple sclerosis, diabetes, obesity, bronchial asthma, acquired
CC   immunodeficiency syndrome (AIDS), Crohn's disease, infectious disease,
CC   anorexia and immune disorders. This sequence represents a human NOVX
CC   polynucleotide of the invention. Note: The sequence data for this patent
XX   is also available from USPTO at seqdata.uspto.gov/sequence.html.
SQ   Sequence 1585 BP; 263 A; 531 C; 486 G; 305 T; 0 U; 0 Other;

Query Match             99.6%;  Score 1520.6;  DB 12;  Length 1585;
Best Local Similarity   99.7%;  Pred. No. 0;
Matches 1523;  Conservative   0;  Mismatches   4;  Indels   0;  Gaps   0;

Qy          1 ATGGACCACCTCGGGGCGTCCCTCTGCCCCAGTCGGCTCGCTCCCTTGTCTCCTGCTCGCT 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         12 ATGGACCACCTCGGGGCGTCCCTCTGCCCCAGTCGGCTCGCTCCCTTGTCTCCTGCTCGCT 71

Qy         61 GGGGCGCCTGGCGCCCCCGCCTAACCTCCGGACCCCAAGTCGAGAGCAAAGCGGCC 120
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         72 GGGGCGCCTGGCGCCCCCGCCTAACCTCCGGACCCCAAGTCGAGAGCAAAGCGGCC 131

Qy        121 TTGCTGCGGCGCCGGGGCCCGAAGAGCTTCGTGCTTCACCGAGCGGTTGGAGGACTTG 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        132 TTGCTGCGGCGCCGGGGCCCGAAGAGCTTCGTGCTTCACCGAGCGGTTGGAGGACTTG 191

Qy        181 GTGTGTTTCTGGGAGGAAGCGGCGGAGCGCTGGGGTGGGCCCGGCCAACTACAGCTTCTCC 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        192 GTGTGTTTCTGGGAGGAAGCGGCGGAGCGCTGGGGTGGGCCCGGCCAACTACAGCTTCTCC 251

Qy        241 TACCAGCTCGAGGATGAGCCATGGAGCTGTGTCGCCTGACCAGGCTCCCACGGCTCGT 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        252 TACCAGCTCGAGGATGAGCCATGGAGCTGTGTCGCCTGACCAGGCTCCCACGGCTCGT 311

Qy        301 GGTGCGGTGGCGCTTCGTGGTGTTCGCTTCGCTGCCTACAGCGCCGCGCTCCGAGCTTCGTGGTGCCCCTA 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        312 GGTGCGGTGGCGCTTCGTGGTGTTCGCTTCGCTGCCTACAGCGCCGCGCTCCGAGCTTCGTGGTGCCCCTA 371

Qy        361 GAGTTGCGGTCACAGCAGCCTCCGGCGTCCGGCGCCCGGATATCACCGTGTCATCCACATCAAT 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        372 GAGTTGCGGTCACAGCAGCCTCCGGCGTCCGGCGCCCGGATATCACCGTGTCATCCACATCAAT 431

Qy        421 GAAGTAGTGCTTCCTAGACGCCCCGTGGGGCGTCGGTGGTGCGCGGTTGGCTGGTGACGAGAGCGGC 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        432 GAAGTAGTGCTTCCTAGACGCCCCGTGGGGCGTCGGTGGTGCGCGGTTGGCTGGTGACGAGAGCGGC 491

Qy        481 CACGTAGTGTTGCGCGTGGCTCCGGCTGAGACACCATGAGCGTCTCACATCCGCTAC 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        492 CACGTAGTGTTGCGCGTGGCTCCGGCTGAGACACCATGAGCGTCTCACATCCGCTAC 551

Qy        541 GAGGTGGACGTCTCGGCCGCAACGGCGGCAGGAGCGGTACAGAGGTGGAGATCCTGGAG 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        552 GCGGTGGACGTCTCGGCCGCAACGGCGGCAGGAGCGGTACAGAGGTGGAAGATCCTGGAG 611

Qy        601 GGCCGCACCGAGTGTGTGCTGAGCAACCTGGGCCGGACGCGCGACACCTTCGCCGTC 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        612 GGCCGCACCGAGTGTGTGCTGAGCAACCTGGGCCGGACGCGCGACACCTTCGCCGTC 671

Qy        661 CGCGGCGTATGGCGTGAGCGCCTTCGGCGGCGCCTTCGGAGCCCTGTG 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        672 CGCGGCGTATGGCGTGAGCGCCTTCGGCGGCGCCTTCGGAGCCGGAGCCCTGTG 731

Qy        721 TCGCTGCTGACGCCTAGGGACGCTGGACCCCGTCTCCCATCCTCGTG 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        732 TCGCTGCTGACGCCTAGGGACGCTGGACCCCGTCTCCCATCCTCGTG 791

Qy        781 GTCATCCTGGTGCTGCTGACCGTGCTCGCGCTGCTCCCACCGCGGGGCTCTGAAGCAG 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        792 GTCATCCTGGTGCTGCTGACCGTGCTCGCGCTGCTCCCACCGCGGGGCTCTGAAGCAG 851

Qy        841 AAGATGCGGCTGGCATCCCGAGCCCAGAGAGCGAGTTTGAAGGCCTCTTCACCACCCAC 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        852 AAGATCTGGCCTGGCATCCCGAGCCCAGAGAGCGAGTTTGAAGGCCTCTTCACCACCCAC 911

Qy        901 AAGGGTAACTTCCAGCTGTGCTGTACCAGAATGATGGCTGCCTGTGGTGGAGCCCTGC 960
```

```
XX
PA   (GEMY )  GENETICS INST INC.
PA   (CHIL-)  CHILDRENS MEDICAL CENT.
PA   (WHED )  WHITEHEAD INST BIOMEDICAL RES.
XX
PI   D'andrea A,   Wong GG,   Jones SS;
XX
DR   WPI;  1994-025409/03.
DR   P-PSDB;  AAR47518.
XX
PT   Recombinant DNA encoding erythropoietin receptor - used to develop prods.
PT   for study, treatment or diagnosis of disorders in which receptor is
PT   dysfunctional.
XX
PS   Claim 7;  Fig 9;  24pp;  English.
XX
CC   Mouse erythroleukaemia (MEL) cells were used to construct a cDNA library.
CC   The cDNA was used to transfect COS-1 cells and these were screened for
CC   radioiodinated erythropoietin (EPO) binding to isolate cDNA encoding the
CC   EPO receptor. This cDNA was used as a probe to screen a human genomic
CC   cDNA library to obtain DNA encoding the human EPO receptor. The cDNA may
CC   be used to study, treat or diagnose disorders in which the EPO receptor
CC   is dysfunctional. See also AAQ53994. (Updated on 25-MAR-2003 to correct
CC   PF field.)
XX
SQ   Sequence 1818 BP;  308 A;  579 C;  571 G;  360 T;  0 U;  0 Other;
```

```
Query Match             99.3%;  Score 1515.8;  DB 2;  Length 1818;
Best Local Similarity   99.5%;  Pred. No. 0;
Matches 1520;  Conservative  0;  Mismatches  7;  Indels  0;  Gaps  0;
```

```
Qy        1 ATGGACCACCTCGGGGCGTCCCTCTGGCCCCCAGTCGGGGTCTCCCTTGTCTTCCTGCTCGCT   60
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      106 ATGGACCACCTCGGGGCGTCCCTCTGGCCCCCAGTCGGGGTCTCCCTTGTCTTCCTGCTCGCT  165

Qy       61 GGGGCCGCGCTGGGCGCCCCGCCCCCGGGGCCTAACCTCCCGACCCCCAAGTTCGAGAGCAAAGCGGCC  120
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      166 GGGGCCGCGCTGGGCGCCCCGCCCCCGGGGCCTAACCTCCCGACCCCCAAGTTCGAGAGCAAAGCGGCC  225

Qy      121 TTGCTGGCGCCGGGGCCGAGAGCTTCTGTGCTTCACGAGCGGTTGGAGGACTTG  180
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      226 TTGCTGGCGCCGGGGCCGAGAGCTTCTGTGCTTCACGAGCGGTTGGAGGACTTG  285

Qy      181 GTGTGTTTCTGGAGGAAGCGGCGGAGCCCAGCGCCGGGGTGGGCCCGGCAACTACAGCTTCTCC  240
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      286 GTGTGTTTCTGGAGGAAGCGGCGGAGCCCAGCGCCGGGGTGGGCCCGGCAACTACAGCTTCTCC  345

Qy      241 TACCAGCTCGAGGATGAGCCATGAGCAGCCCGCCTGCGAAGCTGTGTCGCCTCCCACGCGCTCGT  300
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      346 TACCAGCTCGAGGATGAGCCATGAGCAGCCCGCCTGCGAAGCTGTGTCGCCTGCACCAGGCTCGT  405

Qy      301 GGTGCGGTCGCCTTCTGTGTTCGCTGCTGGTGTTCGCTGCTCACAGCCGACACGTCGTGCCCCTA  360
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      406 GGTGCGGTCGCCTTCTGTGTTCGCTGCTGGTGTTCGCTGCTCACAGCCGACACGTCGTGCCCCTA  465

Qy      361 GAGTTGCGGTCACAGCCTCCGGCGCTCCGGCGATATCCGCGATATCCACATCAAT  420
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      466 GAGTTGCGGTCGACAGCCTCCGGCGCTCCGGCGATATCCGCGATATCCACATGAAT  525

Qy      421 GAAGTAGTGCTCCTAGACGCGCGGGCCCCGTGGGCGGTCGTGGTGGCGCTGACGAGAGCGGGC  480
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      526 GAAGTAGTGCTCCTAGACGCGCGGGCCCCGTGGGCGGTCGTGGTGGCGCTGACGAGAGCGGGC  585

Qy      481 CACCTAGTGTTGCGCTGCTTCCGTTCTGTTCGTTCTGTTCGCTGCTCCAGCTCGAGCCTCCTG  540
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      586 CACCTAGTGTTGCGCTGCTTCCGTTCTGTTCGTTCTGTTCGCTGCTCCAGCTCGAGCCTCCTAC  645

Qy      541 GAGTGGACGTCTCGGCCCGCAACGCGCAGGGAGGCGTACAGAGGGTGGAGATCCTGGAG  600
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      646 GAGTTGGACGTCTCGGCCCGCAACGCGCAGGGAGGCGTACAGAGGGTGGAGATCCTGGAG  705

Qy      601 GGCCGCACCGAGTGTGTGCTGACGAACCTGCGGGCGCCGGCCTACACCTTCGCCGTC  660
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db      912 AAGGGTAACTTCCAGCTGTGGCTGTACCAGAATGATGGCTGCCTGTGGTGGAGCCCCTGC  971
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy      961 ACCCCCTTCACGGAGGAGGACCACCTGCTTCCTGGAAGTCCTTCCCTGGGGG  1020
            ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      972 ACCCCCTTCACGGAGGAGGACCACCTGCTTCCTGGAAGTCCTTCCCTGGGGG  1031
            ||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy     1021 ACGATGCAGGCAGTGAGCCGGGACAGATGAGGGCCCCCTGCTGAGCCAGTGGGC  1080
            ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1032 ACGATGCAGGCAGTGAGCCGGGACAGATGAGGGCCCCCTGCTGAGCCAGTGGGC  1091
            ||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy     1081 AGTGAGCATGCCCAGGATACCTATCTGGTCGTGCGGACAAATGGTTGCTGCCCCCGGAACCCG  1140
            ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1092 AGTGAGCATGCCCAGGATACCTATCTGGTCGTGCTGACAAATGGTTGCTGCCCCCGGAACCCG  1151
            ||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy     1141 CCCAGTGAGGACACTTCCCAGGGCCTGGTGTGGCAGCATGTGGACATAGTGGCCATGGATGAAGGC  1200
            ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1152 CCCAGTGAGGACACTTCCCAGGGCCTGGTGTGGCAGCATGTGGACATAGTGGCCATGGATGAAGGC  1211
            ||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy     1201 TCAGAAGCAGCATCCTCCTGCTCATCTGCTTTGGCCTTGCCTCGAAGCCCAGACCCAGGAGCCTCT  1260
            ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1212 TCAGAAGCAGCATCCTCCTGCTCATCTGCTTTGGCCTTGCCTCGAAGCCCAGACCCAGGAGCCCTCT  1271
            ||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy     1261 GCTGCCAGCTTGAGTACACTATCCTGACCCCCAGCTCCCAGCTCTTGCGTCCATGGACA  1320
            ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1272 GCTGCCAGCTTGAGTACACTATCCTGACCCCCAGCTCCCAGCTCTTGCGTCCATGGACA  1331
            ||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy     1321 CTGTGCCCTGAGCTCGCCCCCTACCCCACCTAAAGTACCTGTACCTTGTGGTATCT  1380
            ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1332 CTGTGCCCTGAGCTCGCCCCCTACCCCACCTAAAGTACCTGTACCTTGTGGTATCT  1391
            ||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy     1381 GACTCTGGCATCTCAACTGACTACAGCTCAGGGGACTCCAGGGAGCCCAAGGGGGCTTA  1440
            ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1392 GACTCTGGCATCTCAACTGACTACAGCTCAGGGGACTCCAGGGAGCCCAAGGGGGCTTA  1451
            ||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy     1441 TCCGATGGCCCCTACTCCAACCCTTATGAGAACAGCCTTATCCCAGCCGCTGCCTCCTG  1500
            ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1452 TCCGATGGCCCCTACTCCAACCCTTATGAGAACAGCCTTATCCCAGCCGCTGCCTCTG  1511
            ||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy     1501 CCCCCAGCTATGTGGCTTGGCTTGCTCTTAG  1527
            ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1512 CCCCCAGCTATGTGGCTTGGCTTGCTCTTAG  1538
```

```
RESULT 11
AAQ53995
ID   AAQ53995 standard;  cDNA;  1818 BP.
XX
AC   AAQ53995;
XX
DT   25-MAR-2003  (revised)
DT   24-JUN-1994  (first entry)
XX
DE   Human EPO receptor cDNA.
XX
KW   Erythropoietin receptor;  recombinant;  murine;  anaemia;  ss.
XX
OS   Homo sapiens.
XX
FH   Key            Location/Qualifiers
FT   CDS            106..1632
FT                  /*tag= a
FT   sig_peptide    106..177
FT                  /*tag= b
FT                  /note= "signal peptide"
XX
PN   US5278065-A.
XX
PD   11-JAN-1994.
XX
PF   25-MAR-1991;  91US-00678877.
XX
PR   03-FEB-1989;  89US-00306503.
```

KW   Truncated human EpoR; erythropoietin receptor; hypersensitive EpoR(t439);
KW   mutant human EpoR; EpoR signalling; cancer; infectious disease; HIV;
KW   sickle cell anaemia; cytostatic; antimicrobial; antiviral;
KW   immunostimulant; anti-anaemic; ss.
XX
OS   Homo sapiens.
XX
FH   Key   Location/Qualifiers
FT   CDS   1..1317
FT   /*tag= a
FT   /product= "Truncated human EpoR(t439)"
XX
PN   WO9967360-A2.
XX
PD   29-DEC-1999.
XX
PF   25-JUN-1999;   99WO-CA000606.
XX
PR   25-JUN-1998;   98CA-02241576.
PR   25-JAN-1999;   99CA-02260332.
XX
PA   (HEMO-) HEMOSOL INC.
XX
PI   Bell D, Matthews KE, Mueller SG;
XX
DR   WPI; 2000-136979/12.
DR   P-PSDB; AAY44622.
XX
PT   Serum free defined medium useful for the efficient culture of stem cells
PT   used for production of hemoglobin.
XX
PS   Example 6; Fig 9; 61pp; English.
XX
CC   The present cDNA sequence encodes truncated human EpoR (erythropoietin
CC   receptor). This was isolated from human umbilical cord blood LDMNC using
CC   standard PCR methods. Transfection of constitutively active EpoR(t439) by
CC   electroporation into a cytokine-dependent cell line supports cell
CC   population expansion in the absence of exogenous cytokines. Mutant human
CC   EpoR is used in treatment of disorders related to inadequate EpoR
CC   signalling. The transfected cells may also used in gene therapy to treat
CC   cancer, infectious diseases (e.g. HIV), sickle cell anaemia, and
CC   conditions related to abnormal expression of erythropoietin
XX
SQ   Sequence 1317 BP; 207 A; 437 C; 422 G; 251 T; 0 U; 0 Other;

Query Match   85.9%; Score 1312.2; DB 3; Length 1317;
Best Local Similarity   99.8%; Pred. No. 2.3e-272;
Matches 1314; Conservative   0; Mismatches   3; Indels   0; Gaps   0;

Qy   1 ATGGACCACCTCGGGCGTCCCTCTGGCCCCAGTCGGCTCCCTTTGTCTCCTCGGCTCCGCT 60
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1 ATGGACCACCTCGGGCGTCCCTCTGGCCCCAGTCGGCTCCCTTTGTCTCCTCGGCTCCGCT 60

Qy   61 GGGGCCGCGCTGGGCGCCCGGGGCCCCGGGCCCTAACCTCCCGGACCCCCCAAGTTCGAGAGCAAAGCGGCC 120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61 GGGGCCGCGCTGGGCGCCCGGGGCCCCGGGCCCTAACCTCCCGGACCCCCCAAGTTCGAGAGCAAAGCGGCC 120

Qy   121 TTGCTGGCGGCCCCGGGCCCGGGGTGCTTCTGTGCTTCACCGAGCGGTTGGAGGACTTG 180
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   121 TTGCTGGCGGCCCCGGGCCCGGGGTGCTTCTGTGCTTCACCGAGCGGTTGGAGGACTTG 180

Qy   181 GTGTGTTCTGGAGGAGCGGCGGCGAGCGGAGCGGCCCGGGCCCCGGGCCAACTACAGCTTCTCC 240
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   181 GTGTGTTCTGGAGGAGCGGCGGCGAGCGGAGCGGCCCGGGCCCCGGGCCAACTACAGCTTCTCC 240

Qy   241 TACCAGTTCGAGGATGAGCCATGGAAGCTGTGTCGCCTGCACCAGGCTCCCACGGCTCGT 300
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   241 TACCAGTTCGAGGATGAGCCATGGAAGCTGTGTCGCCTGCACCAGGCTCCCACGGCTCGT 300

Qy   301 GGTGGCGGTGCGCTTCGTGTTCGCTGCTACAGCCGGACACGGCTTCGTGCCCCTA 360
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   301 GGTGGCGGTGCGCACGCTTCGTGTTCGCTGCTACACAGCCGGACACGGCTTCGTGCCCCCTA 360

Db   706 GGCCGCACCGAGTGTGCTGAGCAACCTGCGGGGCCGGACGCGCTACACCTTCGCCGTC 765
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   661 CGGCGGCGGTATGGCTGAGCCGAGCTTCGGCGGCGTTCGGAGGCGCTGGTGGAGCCTGTG 720
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   766 CGGCGGCGGTATGGCTGAGCCGAGCTTCGGCGGCGTTCGGAGGCGCTGGTGGAGCCTGTG 825

Qy   721 TCGCTGCTGACGACCTAGCGACCTCTGACGCTCTCCCTCATCCTGACGCTCTCCCTCATCCTCGTG 780
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   826 TCGCTGCTGACGCGCCTCAGCGACCTCTGACGCCCCTCACCTCGACGCTCTCCCTCATCCTCGTG 885

Qy   781 GTCATCCTCGTGCTGTGCTGACCGTGCTCGGCGTGCTCGCCGCTCGCTGCTCCTGAAGCAG 840
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   886 GTCATCCTCGTGCTGTGCTGACCGTGCTCGGCGTGCTCGCCGCTCGCCGCTCGCTGCTGAAGCAG 945

Qy   841 AAGATCTGGCCTGGGCATCCCGAGCCCAGAGAGCGAGTTTGAAGGCCTCTTCACCACCCAC 900
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   946 AAGATCTGGCCTGGGCATCCCGAGCCCAGAGAGCGAGTTTGAAGGCCTCTTCACCACCCAC 1005

Qy   901 AAGGGTAACTTCCAGCTGTGGCTGTACCAGAATGGCTGCCTGTGGTGGAGCCCCCTGC 960
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1006 AAGGGTAACTTCCAGCTGTGGCTGTACCAGAATGGCTGCCTGTGGTGGAGCCCCCTGC 1065

Qy   961 ACCCCCTTCACGGAGGACCACCTTGCTTCCCTGGAAGTCCTCTTCAGAGCGGCTGCTGGGGG 1020
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1066 ACCCCCTTGACGGAGGACCACCTTGCTTCCCTGGAAGTCCTCTTCAGAGCGGCTGCTGGGGG 1125

Qy   1021 ACGATGCAGGCAGTGGAGCCGGGGACAGATGATGAGGGCCCCCTGCTCGAGCCAGTGGGC 1080
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1126 ACGATGCAGGCAGTGGAGCCGGGGACAGATGATGAGGGCCCCCTGCTCGAGCCAGTGGGC 1185

Qy   1081 AGTGAGCATGCCCAGGATACCTATCTGGTGCTGGACAAATGGTTGCTGCCCCGGAACCCG 1140
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1186 AGTGAGCATGCCCAGGATACCTATCTGGTGCTGGACAAATGGTTGCTGCCCCGGAACCCG 1245

Qy   1141 CCCAGTGAGGACTTCCCAGGGGCCTGGTGGCAGTTGTGGACATAGTGGACATGATGAAGGC 1200
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1246 CCCAGTGAGGACTTCCCAGGGGCCTGGTGGCAGTTGTGGAGATAGTGGACATGATGAAGGC 1305

Qy   1201 TCAGAAGCATCCTCCTGCTCATCGTTTGGCCTTGGCCTCGAAGCCCAGAGAGGGAGCCTCT 1260
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1306 TCAGAAGCATCCTCCTGCTCATCGTCTTTGGCCTTGGCCTCGAAGCCCAGAGAGGGAGCCTCT 1365

Qy   1261 GCTGCCAGCTTTGAGTAGTACACTATCCTGACACCCCAGCTCTTGCGTCCATGGACA 1320
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1366 GCTGCCAGCTTTGAGTAGTACACTATCCTGACACCCCAGCTCTTGCGTCCATGGACA 1425

Qy   1321 CTGTGCCCTGAGCTGCCCCTCACCCCACCCCCTAAAGTACCTGTACCTTGTGGTATCT 1380
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1426 CTGTGTTTCTGGAGGAAGCGGTCGGCGTGTTCGCTGCTACACCTGTACCTTGTGGTATGT 1485

Qy   1381 GACTCTGGCATCTCAACTGACTACAGCTCAGGGGACTCCAGGGGGAGCCCAAGGGGCTTA 1440
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1486 GACTCTGGCATGTCAACTGACTACAGCTCAGGGGACTCCAGGGGGAGCCCAAGGGGCTTA 1545

Qy   1441 TCCGATGGCCCCTACTCCAACCCTTATGAGAACAGCCCTTATCCCCAGCCGCTGAGCCTCTG 1500
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1546 TCCGATGGCCCCTACTCCAACCCTTATGAGAACAGCCCTTATCCCCAGCCGCTGAGCCTCTG 1605

Qy   1501 CCCCCCAGCTATGTGGCTTGCTCTTAG 1527
           |||||||||||||||||||||||||
Db   1606 CCCCCCAGCTATGTGGCTTGCTCTTAG 1632

RESULT 12
AAZ49634
ID AAZ49634 standard; cDNA; 1317 BP.
XX
AC AAZ49634;
XX
DT 07-APR-2000 (first entry)
XX
DE Truncated human EpoR(t439) cDNA.
XX

```
DE   Mutant R154C human EpoR(t439) cDNA.
XX
KW   Truncated human EpoR; erythropoietin receptor; mutant human EpoR;
KW   EpoR signalling; cancer; infectious disease; HIV, sickle cell anaemia;
KW   cytostatic; antimicrobial; antiviral; immunostimulant; anti-anaemic; ss.
XX
OS   Homo sapiens.
XX
FH   Key            Location/Qualifiers
FT   CDS            1..1317
FT                  /*tag= a
FT                  /product= "R154C human EpoR(t439)"
XX
PN   WO9967360-A2.
XX
PD   29-DEC-1999.
XX
PF   25-JUN-1999;  99WO-CA000606.
XX
PR   25-JUN-1998;  98CA-02241576.
PR   25-JAN-1999;  99CA-02260332.
XX
PA   (HEMO-) HEMOSOL INC.
XX
PI   Bell D, Matthews KE, Mueller SG;
XX
DR   WPI; 2000-136979/12.
DR   P-PSDB; AAY44623.
XX
PT   Serum free defined medium useful for the efficient culture of stem cells
PT   used for production of hemoglobin.
XX
PS   Example 6; Fig 10; 61pp; English.
XX
CC   The present cDNA sequence encodes mutant R154C truncated (at 439 residue)
CC   human EpoR (erythropoietin receptor). This was isolated from human
CC   umbilical cord blood LDMNC using standard PCR methods. Transfection of
CC   constitutively active mutant EpoR (t439) by electroporation into a
CC   cytokine-dependent cell line supports cell population expansion in the
CC   absence of exogenous cytokines. Mutant human EpoR is used in treatment of
CC   disorders related to inadequate EpoR signalling. The transfected cells
CC   may also used in gene therapy to treat cancer, infectious diseases (e.g.
CC   HIV), sickle cell anaemia, and conditions related to abnormal expression
CC   of erythropoietin
XX
SQ   Sequence 1317 BP; 208 A; 437 C; 420 G; 252 T; 0 U; 0 Other;

Query Match             85.6%;  Score 1307.4;  DB 3;  Length 1317;
Best Local Similarity   99.5%;  Pred. No. 2.4e-271;
Matches 1311;  Conservative  0;  Mismatches  6;  Indels  0;  Gaps  0;
```

```
Qy     1 ATGGACCACCTCGGGGCGTCCCTCTGGCCCAGTCGGCTCGGCTCCCTTGTCTCCTGCTCGCT     60
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1 ATGGACCACCTCGGGGCGTCCCTCTGGCCCAGTCGGCTCGGCTCCCTTGTCTCCTGCTCGCT     60

Qy    61 GGGGCCGCCGCCTGGGCGCCCCGCCCCCGGGCTCCCGACCCCAAGTTCGAGAGCAAAGCGGCC    120
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    61 GGGGCCGCCGCCTGGGCGCCCCGCCCCCGGGCTCCCGACCCCAAGTTCGAGAGCAAAGCGGCC    120

Qy   121 TTGCTGGCGCCCGGGGGCGGCCCCGGGGGCCGAAGAGCTTCTGTGCTTCACGGAGCGGTTGGAGGACTTG  180
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   121 TTGCTGGCGCCCGGGGGCGGCCCCGGGGGCCGAAGAGCTTCTGTGCTTCACGGAGCGGTTGGAGGACTTG  180

Qy   181 GTGTGTTCTGGAGGAAGCGGCGAGCGGCGGCAACTACAGCTTCTCC    240
         |||||||||||||||||||||||||||||||||||||||||||||
Db   181 GTGTGTTCTGGAGGAAGCGGCGAGCGGCGGCAACTACAGCTTCTCC    240

Qy   241 TACCAGCTTCGAGGATGAGCCATGGAAGCGTGTCGCCTCCCACGGCTCGT    300
         |||||||||||||||||||||||||||||||||||||||||||||||||
Db   241 TACCAGCTTCGAGGATGAGCCATGGAAGCGTGTCGCCTCCCACGGCTCGT    300

Qy   301 GGTGCGGTGCGCTTCGGTGTTCGCTGCGTCACGCCGACACGTCGAGCTTCGTGCCCCCTA    360
         |||| ||||||||||||||||||||||||||||||
```

```
Qy   361 GAGTTGCGCGGTCACAGCAGCCTCCGGCGCGGATATCACCGTGTCATCCACATCAAT    420
         ||||||||||||||||||||||||||||||||||
Db   361 GAGTTGCGCGGTCACAGCAGCCTCCGGCGCTCCGGGCTGGTCATCACCGTGTCATCCACATCAAT    420

Qy   421 GAAGTAGTGCTTCCTAGACGCCCCCGTGGGCGCTGGTGGCGCGGTTGGCTGACGAGAGCGGC    480
         |||||||||||||||||||||||||||||||||||||||||||||||||
Db   421 GAAGTAGTGCTTCCTAGACGCCCCCGTGGGCGCTGGTGGCGCGGTTGGCTGACGAGAGCGGC    480

Qy   481 CACGTAGTGTTGCGCTGGCTGGCCTGAGACACCCATGAGACGTCTCACATCCGCTAC    540
         |||||||||||||||||||||||||||||||||||||||||||||||||||
Db   481 CACGTAGTGTTGCGCTGGCTGGCCTGAGACACCCATGAGACGTCTCACATCCGCTAC    540

Qy   541 GAGGTGGACGTCTCGGCCGGCAACGGCGCAGGGAGCGGTACAGAGGGTGGAGATCCTGGAG    600
         |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   541 GGCGTGGACGTCTCGGCCGGCAACGGCGCAGGGAGCGGTACAGAGGGTGGAGATCCTGGAG    600

Qy   601 GGCCGCACCGAGTGTGTGCTGAGCAACCTGCCGGACGCGCTACACCTTCGCCGTC    660
         |||||||||||||||||||||||||||||||||||||||||||||||||||
Db   601 GGCCGCACCGAGTGTGTGCTGAGCAACCTGCCGGACGCGCTACACCTTCGCCGTC    660

Qy   661 CGCGCGGTATGGCTGAGCCGAGCTTCGGCGGCGCTTCGGCGGCGCCGTCGTG    720
         ||||||||||||||||||||||||||||||||||||||||||||||||||
Db   661 CGCGCGGTATGGCTGAGCCGAGCTTCGGCGGCGCTTCGGCGGCGCCGTCGTG    720

Qy   721 TCGCTGCTGACGCCTAGCGACCTGGACCCCTCATCCTGACGCTCTCCCTCATCCTCGTG    780
         ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   721 TCGCTGCTGACGCCTAGCGACCTGGACCCCTCATCCTGACGCTCTCCCTCATCCTCGTG    780

Qy   781 GTCATCCTGGTGCTGCTGACCGTGCTCGCGCTGCTCTCCCACCGCGGGCTCTGAAGCAG    840
         ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   781 GTCATCCTGGTGCTGCTGACCGTGCTCGCGCTGCTCTCCCACCGCGGGCTCTGAAGCAG    840

Qy   841 AAGATCTGGCTGGCATCCCGAGCCCAGAGAGCGAGTTTGAAGGCCTCTTCACCACCCAC    900
         |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   841 AAGATCTGGCTGGCATCCCGAGCCCAGAGAGCGAGTTTGAAGGCCTCTTCACCACCCAC    900

Qy   901 AAGGGTAACTTCCAGCTGGCTGGCTGTACCAGAATGGCTGGCTGTGTGGTGGAGCCCCTGC    960
         ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   901 AAGGGTAACTTCCAGCTGGCTGTACCAGAATGATGGCTGGCTGTGTGGTGGAGCCCCTGC    960

Qy   961 ACCCCCTTCACGGAGGAGCCACCTGCTTCCTCTTCCTGGAAGTCCTCAGAGCCGTCGTGGGGG   1020
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   961 ACCCCCTTCACGGAGGACCACCTGCTTCCTCTTCCTGGAAGTCCTCAGAGCCGTCGTGGGGG   1020

Qy  1021 ACGATGCAGCAGTGGAGCCGGGGACAGATGATGAGGGCCCCCCTGCTGGAGCCAGTGGGC   1080
         ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1021 ACGATGCAGCAGTGGAGCCGGGGACAGATGATGAGGGCCCCCCTGCTGGAGCCAGTGGGC   1080

Qy  1081 AGTGAGCATGCCCAGGATACCTATCTGGTTGCTGACAAATGGTTGCTGCCCCCGGAACCCG   1140
         |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1081 AGTGAGCATGCCCAGGATACCTATCTGGTTGCTGACAAATGGTTGCTGCCCCCGGAACCCG   1140

Qy  1141 CCCAGTGAGGACCTTGGTGCAGTGGACATAGTGGCAGATTGGGACATAGTGGCAGATGAAGGC   1200
         ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1141 CCCAGTGAGGACCTTGGTGCAGTGGACATAGTGGCAGATTGGGACATAGTGGCAGATGAAGGC   1200

Qy  1201 TCAGAAGCATCCTCTGCTTGGCCTCGAAGGCCAGCCCAGAGGAGCCTCT   1260
         ||||||||||||||||||||||||||||||||||||||||||||||
Db  1201 TCAGAAGCATCCTCTGCTTGGCCTCGAAGGCCAGCCCAGAGGAGCCTCT   1260

Qy  1261 GCTGCCAGCTTTGAGTACACTATCCTGACCACTATCCTGACGCTCCCAGCTCTTGCGTCCATGG   1317
         ||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1261 GCTGCCAGCTTTGAGTACACTATCCTGACCACTATCCTGACGCTCCCAGCTCTTGCGTCTTCCATAG   1317
```

```
RESULT 13
AAZ49636
ID   AAZ49636 standard; cDNA; 1317 BP.
XX
AC   AAZ49636;
XX
DT   07-APR-2000  (first entry)
XX
```

DT 29-JAN-2004 (first entry).
XX
DE Human NOV15a cDNA - SEQ ID 47.
XX
KW NOVX; antidiabetic; anorectic; cardiant; hypotensive;
KW antiarteriosclerotic; virucide; antibacterial; fungicide; protozoacide;
KW nootropic; neuroprotective; antiparkinsonian; anticonvulsant;
KW osteopathic; antiarthritic; antiinflammatory; dermatological;
KW antiasthmatic; antilipaemic; metabolic; diabetes; obesity; infectious;
KW anorexia; cancer; cardiovascular; hypertension; atherosclerosis;
KW neurodegenerative; Alzheimer's disease; Parkinson's; epilepsy; immune;
KW osteoarthritis; haemopoietic; inflammatory skin; asthma; dyslipidaemia;
KW neurogenesis; cell differentiation; proliferation; haemopoiesis;
KW wound healing; angiogenesis; gene therapy; chromosome mapping;
KW tissue typing; human; ss; gene; NOV.
XX
OS Homo sapiens.
XX
PN WO200304330-A2.
XX
PD 15-MAY-2003.
XX
PF 05-NOV-2002; 2002WO-US035536.
XX
PR 05-NOV-2001; 2001US-0338626P.
PR 05-DEC-2001; 2001US-0336600P.
PR 07-DEC-2001; 2001US-0338628P.
PR 12-DEC-2001; 2001US-0341346P.
PR 17-DEC-2001; 2001US-0341477P.
PR 17-DEC-2001; 2001US-0341540P.
PR 20-DEC-2001; 2001US-0342592P.
PR 27-DEC-2001; 2001US-0344297P.
PR 31-DEC-2001; 2001US-0044903P.
PR 17-APR-2002; 2002US-0373288P.
PR 15-MAY-2002; 2002US-0380981P.
PR 17-MAY-2002; 2002US-0081499P.
PR 28-MAY-2002; 2002US-0383534P.
PR 28-MAY-2002; 2002US-0383744P.
PR 29-MAY-2002; 2002US-0383829P.
PR 29-MAY-2002; 2002US-0384024P.
PR 07-AUG-2002; 2002US-0401788P.
PR 26-AUG-2002; 2002US-0406355P.
PR 31-OCT-2002; 2002US-00287971.
XX
PA (CURA-) CURAGEN CORP.
XX
PI Alsobrook JP, Alvarez E, Anderson DW, Baron M, Boldog FL;
PI Burgess CE, Casman SU, Chapoval A, Dhanabal M, Edinger SR, Eisen A;
PI Ellerman K, Ettenberg S, Gangolli EA, Gerlach VL, Gorman L;
PI Grosse WM, Guo X, Hackett C, Ji W, Kekuda R, Khramtsov NV;
PI Lepley DM, Li L, Macdougall JR, Malyankar UM, Mazur A, Mcqueeney K;
PI Mezes PS, Miller CE, Millet I, Mishra VS, Padigaru M, Patturajan M;
PI Pena CEA, Peyman JA, Rastelli L, Rieger DK, Shenoy SG, Shimkets RA;
PI Smithson G, Starling G, Spytek KA, Stone DJ, Tchernev VT, Twomlow N;
PI Vernet CAM, Zerhusen BD, Zhong M;
XX
DR WPI; 2003-441555/41.
DR P-PSDB; ADE28671.
XX
PT New isolated NOVX polypeptides and polynucleotides, useful for
PT preventing, diagnosing or treating NOVX-associated disorders, e.g.
PT osteoarthritis, obesity, atherosclerosis, cancer, Parkinson's disease,
PT asthma, or infections.
XX
PS Claim 20; SEQ ID NO 47; 447pp; English.
XX
CC The invention relates to a novel isolated NOVX polypeptide. The
CC polypeptide of the invention demonstrates, antidiabetic, anorectic,
CC cardiant, hypotensive, antiarteriosclerotic, virucide, antibacterial,
CC fungicide, protozoacide, nootropic, neuroprotective, antiparkinsonian,
CC anticonvulsant, osteopathic, antiarthritic, antiinflammatory,
CC dermatological, antiasthmatic and antilipaemic activities. The
CC polypeptides, nucleic acid molecules and antibodies may be useful for

Db 301 GGTGCCGTACGCTTCTGGTGTTCGCTGCCTACAGCCGACACGTCGAGCTTCGTGCCCCTA 360

Qy 361 GAGTTGCGCGGTCACAGCAGCCTCCGGCGCTCCGGCGATATCACCGTGTCATCCACATCAAT 420
    ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 361 GAGTTGCGCGGTCACAGCAGCCTCCGGCGCTCCGGCGATATCACCGTGTCATCCACATCAAT 420

Qy 421 GAAGTAGTGCTCCTAGACGCGCCCCGTGGGCGTGGTGCGCGGTTGGCTGACGAGAGCGGC 480
    ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 421 GAAGTAGTGCTCCTAGACGCGCCCCGTGGGCGTGGTGCGCGGTTGGCTGACGAGAGCGGC 480

Qy 481 CACGTAGTTGTTGCGCGTCCGGCTCCGCGCCTGAGACACCCATGACGTCTCACATCCGTAC 540
    ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 481 CACGTAGTTGTTGCGCGTCCGGCTCCGCGCCTGAGACACCCATGACGTCTCACATCCGTAC 540

Qy 541 GAGGTGGACGTCTCGGCCGCGCAACGGCCGCAGGGAGCGTACAGAGGGTGGAGATCCTGAG 600
    ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 541 GAGGTGGACGTCTCGGCCGCGCAACGGCCGCAGGGAGCGTACAGAGGGTGGAGATCCTGAG 600

Qy 601 GGCGCCACCGAGTGTGTGCTGAGCAACCTGCGGGGCGGGACGCGCTACACCTTCGCCGTC 660
    ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 601 GGCGCCACCGAGTGTGTGCTGAGCAACCTGCGGGGCGGGACGCGCTACACCTTCGCCGTC 660

Qy 661 CGCGCGCGTATGGCTGAGCCGAGCTTCGGCGGCGCTTCGGAGCGCTGTGGAGCCTGTG 720
    ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 661 CGCGCGCGTATGGCTGAGCCGAGCTTCGGCGGCGCTTCGGAGCGCTGTGGAGCCTGTG 720

Qy 721 TCGCTGCTGACGGCTAGGGACGCTGGACCCCCTCATCCTGACGCTCTCCCTCATCCTCGTG 780
    ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 721 TCGCTGCTGACGGCTAGGGACGCTGGACCCCCTCATCCTGACGCTCTCCCTCATCCTCGTG 780

Qy 781 GTCATCCTGGTGCTGCTGACCGTGCTCGCCGTGCTCTCCCACCGCGGGCTCTGAAGCAG 840
    ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 781 GTCATCCTGGTGCTGCTGACCGTGCTCGCCGTGCTCTCCCACCGCGGGCTCTGAAGCAG 840

Qy 841 AAGATCTGGCCTGGCATCCCGAGAGAGCCCAGAGCGAGTTTGAAGGCCTTCTTCACCACCAC 900
    ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 841 AAGATCTGGCCTGGCATCCCGAGAGAGCCCAGAGCGAGTTTGAAGGCCTTCTTCACCACCAC 900

Qy 901 AAGGGTAACTTCCAGCTGTGGCTGTACCAGAATGATGGCTGCCTGTGTGGAGCCCCTGC 960
    ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 901 AAGGGTAACTTCCAGCTGTGGCTGTACCAGAATGATGGCTGCCTGTGTGGAGCCCCTGC 960

Qy 961 ACCCCCTTCACGGAGGACCCACCTGCTTCCCTGGAAGTCCTCTCCTGAAGTGCTGCTGGGGG 1020
    ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 961 ACCCCCTTCACGGAGGACCCACCTGCTTCCCTGGAAGTCCTCTCCTGAAGTGCTGCTGGGGG 1020

Qy 1021 ACGATGCAGGCAGTGGACACGCGGGACGAGATGATGAGGGCCCTGCTCGGAGCCAGTGGGC 1080
    ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1021 ACGATGCAGGCAGTGGAGCGCGGGGACGAGATGATGAGGGCCCTGCTCGGAGCCAGTGGGC 1080

Qy 1081 AGTGAGCATGCCCAGGATACCTATCTGGTGCTGACAAATGTTGCTGCCCCGGAACCCG 1140
    ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1081 AGTGAGCATGCCCAGGATACCTATCTGGTGCTGACAAATGTTGCTGCCCCGGAACCCG 1140

Qy 1141 CCCAGTGAGGACCTCCCCAGGGCCTGGTGCAGTGTGGACATAGTGGCCATGATGAAGGC 1200
    ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1141 CCCAGTGAGGACCTCCCCAGGGCCTGGTGCAGTGTGGACATAGTGGCCATGATGAAGGC 1200

Qy 1201 TCAGAAGCATCTCCTGCTCATCCGCTTTGGCCTTTGGCCCTGCCGAGCCCAGCCCAGAGGAGGCCTCT 1260
    ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1201 TCAGAAGCATCTCCTGCTCATCCGCTTTGGCCTTTGGCCCTGCCGAGCCCAGCCCAGAGGAGGCCTCT 1260

Qy 1261 GCTGCCAGCTTTGAGTACACTATCCTGACCCCCAGCTCTTGCTGCTCCATGG 1317
    |||||||||||||||||||||||||||||||||||||||||||||||||
Db 1261 GCTGCCAGCTTTGAGTACACTATCCTGACCCCCAGCTCTTGCTGCTCCATAG 1317

RESULT 14
ADE28670
ID ADE28670 standard; cDNA; 1733 BP.
XX
AC ADE28670;
XX

```
CC    treating or diagnosing diseases including metabolic disorders such as
CC    diabetes and obesity, infectious diseases, anorexia, cancer,
CC    cardiovascular diseases including hypertension and atherosclerosis,
CC    neurodegenerative disorders such as Alzheimer's disease, Parkinson's
CC    disease and epilepsy, immune disorders e.g. osteoarthritis, haemopoietic
CC    disorders, inflammatory skin disorders, asthma and dyslipidaemia.
CC    Furthermore, the nucleic acids and polypeptides may also be used to
CC    identify molecules that modulate or inhibit neurogenesis, cell
CC    differentiation and proliferation, haemopoiesis, wound healing and
CC    angiogenesis, as well as in gene therapy. Finally, the nucleic acids may
CC    be used as hybridisation probes, in chromosome mapping, tissue typing,
CC    preventive medicine and pharmacogenomics. The current sequence is that of
CC    the human NOV cDNA of the invention.
XX
SQ    Sequence 1733 BP; 291 A; 558 C; 546 G; 338 T; 0 U; 0 Other;

Query Match             79.2%;  Score 1209; DB 10; Length 1733;
Best Local Similarity   89.8%;  Pred. No. 3.6e-250;
Matches 1372;  Conservative   0;  Mismatches  5;  Indels 150;  Gaps  1;

Qy      1 ATGACCACCTCGGGGCGTCCCTCGGCTCGGCTCCCTTGTCTCCTGCTGCT 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||
Db    145 ATGACCACCTCGGGGCGTCCCTCGGCTCGGCTCCCTTGTCTCCTGCTGCT 204

Qy     61 GGGGCCGCCTCGGGGCGCCCCCCGGACCCCAAGTTCGAGAGCAAAGCGGCC 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||
Db    205 GGGGCCGCCTCGGGGCGCCCCCCGGACCCCAAGTTCGAGAGCAAAGCGGCC 264

Qy    121 TTGCTGGGCGGCCCCGGGGGGCCGAAGAGCTTCTGTGCTTCACCGAGCGGTTGGAGGACTTG 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||
Db    265 TTGCTGGGCGGCCCCGGGGGGCCGAAGAGCTTCTGTGCTTCACCGAGCGGTTGGAGGACTTG 324

Qy    181 GTGTGTTTCTGGGAGGAAGCGGCGAGCGGCGGCAACTACAGCTTCTCC 240
          ||||||||||||||||||||||||||||||||||||||||||||||||
Db    325 GTGTGTTTCTGGGAGGAAGCGGCGAGCGGCGGCAACTACAGCTTCTCC 384

Qy    241 TACCAGCTCGAGGATGAGCCATGGAAGTCGTGCCTGCACCAGGCTCGT 300
          ||||||||||||||||||||||||||||||||||||||||||||||||
Db    385 TACCAGCTCGAGGATGAGCCATGGAAGTCGTGCCTGCACCAGGCTCGT 444

Qy    301 GGTGCGGGTGCGCTTCTGGTGTTCGCTGCTACAGCCGACACGTCGAGCTTCGTGCCCCTA 360
          ||||||||||||||||||||||||||||||||||||||||||||||||
Db    445 GGTGCGGGTGCGCTTCTGGTGTTCGCTGCTACAGCCGACACGTCGAGCTTCGTGCCCCTA 504

Qy    361 GAGTTGCGGCTCACAGCAGCCTCCGGCGCTCCGGCGATATCACCGTGTCATCCACATCAAT 420
          ||||||||||||||||||||||||||||||||||||||||||||||||
Db    505 GAGTTGCGGCTCACAGCAGCCTCCGGCGCTCCGGCGATATCACCGTGTCATCCACATCAAT 564

Qy    421 GAAGTAGTGCTCCTAGACGCGCCCGTGGGCGGTTGGCTGACGAGAGCGGC 480
          ||||||||||||||||||||||||||||||||||||||||||||||||
Db    565 GAAGTAGTGCTCCTAGACGCGCCCGTGGGCGGTTGGCTGACGAGAGCGGC 624

Qy    481 CACGTAGTGTTGCGCTGGCTCCCCGCCGCTGACACACCCATGACGTCTCACATCCGTAC 540
          ||||||||||||||||||||||||||||||||||||||||||||||||
Db    625 CACGTAGTGTTGCGCTGGCTCCCCGCCGCTGACACACCCATGACGTCTCACATCCGTAC 684

Qy    541 GAGTGGACGTCTCGGCCGCGGCAACGGCGGAGGCGTACAGAGGGTGGAGATCCTGGAG 600
          ||||||||||||||||||||||||||||||||||||||||||||||||
Db    685 GAGTGGACGTCTCGGCCGCGGCAACGGCGGAGGCGTACAGAGGGTGGAGATCCTGGAG 744

Qy    601 GGGCCGCCAGAGTGTGCTGCGAGCAACCTGCGGCGGCGACGCGGGACCTACACCTTTCGCCGTC 660
          ||||||||||||||||||||||||||||||||||||||||||||||||
Db    745 GGGCCGCCAGAGTGTGCTGCGAGCAACCTGCGGCGGCGACGCGGGACCTACACCTTTCGCCGTC 804

Qy    661 CCGCGGCGTATGGCTGAGCCGAGCTTCGGCGGCTTCGGCGGCGCCTGTCGGAGCCGTGTG 720
          ||||||||||||||||||||||||||||||||||||||||||||||||
Db    805 CCGCGGCGTATGGCTGAGCCGAGCTTCGGCGGCTTCGGCGGCGCCTGTCGGAGCCGTGTG 864

Qy    721 TCGCTGCTGACGCGCTAGGACCTGGACCCCCATCCTCGACGCGCTCCCATCCTCGTG 780
          ||||||||||||||||||||||||||||||||||||||||||||||||
Db    865 TCGCTGCTGACGCGCTAGGACCTGGACCCCCATCCTCGACGCGCTCCCATCCTCGTG 924

Qy    781 GTCATCCGTGCTGCTGACCGGTCGCTCCCCACCCCGGGGCTCTGAAGCAG 840
          ||||||||||||||||||||||||||||||||||||||||||||||||

Db    925 GTCATCCTGGTGCTGCTGACCGGTCGCTCCCCACCCCGGGGCTCTGAAGCAG 984

Qy    841 AAGATCTGGCTCGCCATCCGAGCCCAGAGACGGAGTTGAAGGCCTCTTCACCACCCAC 900
          ||||||||||||||||||||||||||||||||||||||||||||||||
Db    985 AAGATCTGGCTCGCCATCCGAGCCCAGAGACGGAGTTGAAGGCCTCTTCACCACCCAC 1044

Qy    901 AAGGGTAACTTCCAGCTGGCTGTACCAGAATGATGGCTGCTGTGGTGGAGCCCCTGC 960
          ||||||||||||||||||||||||||||||||||||||||||||||||
Db   1045 AAGGGTAACTTCCAGCTGGCTGTACCAGAATGATGGCTGCTGTGGTGGAGCCCCTGC 1104

Qy    961 ACCCCCTCACGGAGGACGACCACCTGCTTCCCTGGAAGTCCTCTCAGAGCGCTGTGGGGG 1020
          ||||||||||||||||||||||||||||||||||||||||||||||||
Db   1105 ACCCCCTTCACGGAGGACGACCACCTGCTTCCCTGGAAGTCCTCTCAGAGCGCTGTGGGGG 1164

Qy   1021 ACGATGAGGCAGTGGAGCCGGGACAGATGATGAGGGCCCCTGCTGTGGAGCCAGTGGGC 1080
          ||||||||||||||||||||||||||||||||||||||||||||||||
Db   1165 ACGATGAGGCAGTGGAGCCGGGACAGATGATGAGGGCCCCTGCTGTGGAGCCAGTGGGC 1224

Qy   1081 AGTGAGCATGCCCAGGATACCTATCCTGCTGGACAAATGGTTGCTGCCCCCGGAACCCG 1140
          ||||||||||||||||||||||||||||||||||||||||||||||||
Db   1225 AGTGAGCATGCCCAGGATACCTATCCTGCTGGACAAATGGTTGCTGCCCCCGGAACCCG 1284

Qy   1141 CCCAGTGAGGACCTCCCAGGGCCTCCCAGGG 1200
          |||||||||||||||||||||||||||||||
Db   1285 CCCAGTGAGGACCTCCCAGGG----------- 1305

Qy   1201 TCAGAAGCATCCTCCTGCTCATCTGCTTTGGCCTCGAAGCCCAGAGCCCAGAGGGAGCCTCT 1260
          |||||||||||||||||||||||||||||||
Db   1306 ------------------------------- 1305

Qy   1261 GCTGCCAGCTTTGAGTACACTATCCTGACCCCCAGCTCTCCCAGTCTTCCGCGTCCATGGACA 1320
          |||||||||||||||||||||||||||||||          |||||||
Db   1306 ------------------------------CCATGGGCA 1314

Qy   1321 CTGTGCCCTGAGCTGCCCCACCCACCCTAAAGTACCTGTACTTGTGGTATCT 1380
          ||||||||||||||||||||||||||||||||||||||||||||||||
Db   1315 CTGTGCCCTGAGCTGCCCCACCCACCCTAAAGTACCTGTACTTGTGGTATCT 1374

Qy   1381 GACTCTGGCATCTCAACTGACTACAGCTCAGGGGACTCCCAGGGAGCCCAAGGGGGCTTA 1440
          ||||||||||||||||||||||||||||||||||||||||||||||||
Db   1375 GACTCTGGCATCTCAACTGACTACAGCTCAGGGGACTCCCAGGGAGCCCAAGGGGGCTTA 1434

Qy   1441 TCCGATGGCCCCTCCAACCCTTATGAGAACAGCTTATCCCAGCCGCTGCGCCTCTG 1500
          ||||||||||||||||||||||||||||||||||||||||||||||||
Db   1435 TCCGATGGCCCCTACTCCAACCCTTATGAGAACAGCCCTATCCCAGCCGCTGCGCCTCTG 1494

Qy   1501 CCCCCAGCTATGTGGCTTGCTCTTAG 1527
          ||||||||||||||||||||||||||
Db   1495 CCCCCAGCTATGTGGCTTGCTCTTAG 1521
```

RESULT 15
ADE28672
```
ID   ADE28672 standard; cDNA; 1733 BP.
XX
AC   ADE28672;
XX
DT   29-JAN-2004  (first entry)
XX
DE   Human NOV15b cDNA - SEQ ID 49.
XX
KW   NOVX; antidiabetic; anorectic; cardiant; hypotensive;
KW   antiarteriosclerotic; virucide; antibacterial; fungicide; protozoacide;
KW   nootropic; neuroprotective; antiparkinsonian; anticonvulsant;
KW   osteopathic; antiarthritic; antiinflammatory; dermatological;
KW   antiasthmatic; antilipaemic; metabolic; diabetes; obesity; infectious;
KW   anorexia; cancer; cardiovascular; hypertension; atherosclerosis;
KW   neurodegenerative; Alzheimer's disease; Parkinson's; epilepsy; immune;
KW   osteoarthritis; haemopoietic; inflammatory skin; asthma; dyslipidaemia;
KW   neurogenesis; cell differentiation; proliferation; haemopoiesis;
KW   wound healing; angiogenesis; gene therapy; chromosome mapping;
KW   tissue typing; human; ss; gene; NOV.
```

```
XX
OS   Homo sapiens.
XX
PN   WO200304030-A2.
XX
PD   15-MAY-2003.
XX
PF   05-NOV-2002; 2002WO-US035536.
XX
PR   05-NOV-2001; 2001US-0338626P.
PR   05-DEC-2001; 2001US-0336600P.
PR   07-DEC-2001; 2001US-0338285P.
PR   12-DEC-2001; 2001US-0341346P.
PR   17-DEC-2001; 2001US-0341477P.
PR   17-DEC-2001; 2001US-0341540P.
PR   20-DEC-2001; 2001US-0342592P.
PR   27-DEC-2001; 2001US-0344297P.
PR   31-DEC-2001; 2001US-0344903P.
PR   17-APR-2002; 2002US-0373288P.
PR   15-MAY-2002; 2002US-0380981P.
PR   17-MAY-2002; 2002US-0381495P.
PR   28-MAY-2002; 2002US-0383534P.
PR   28-MAY-2002; 2002US-0383744P.
PR   29-MAY-2002; 2002US-0383829P.
PR   29-MAY-2002; 2002US-0384024P.
PR   07-AUG-2002; 2002US-0401788P.
PR   26-AUG-2002; 2002US-0406353P.
PR   31-OCT-2002; 2002US-00287971.
XX
PA   (CURA-) CURAGEN CORP.
XX
PI   Alsobrook JP, Alvarez E, Anderson DW, Baron M, Boldog FL;
PI   Burgess CE, Casman SJ, Chapoval A, Dhanabal M, Edinger SR, Eisen A;
PI   Ellerman K, Ettenberg S, Gangolli EA, Gerlach VL, Gorman L;
PI   Grosse WM, Guo X, Hackett C, Ji W, Kekuda R, Khramtsov NV;
PI   Lepley DM, Li L, Macdougall JR, Malyankar UM, Mazur A, Mcqueeney K;
PI   Mezes PS, Miller CE, Millet I, Mishra VS, Padigaru M, Patturajan M;
PI   Pena CEA, Peyman JA, Rastelli L, Rieger DK, Shenoy SG, Shimkets RA;
PI   Smithson G, Starling G, Spytek KA, Stone DJ, Tchernev VT, Twomlow N;
PI   Vernet CAM, Zerhusen BD, Zhong M;
XX
DR   WPI; 2003-441555/41.
DR   P-PSDB; ADE28673.
XX
PT   New isolated NOVX polypeptides and polynucleotides, useful for
PT   preventing, diagnosing or treating NOVX-associated disorders, e.g.
PT   osteoarthritis, obesity, atherosclerosis, cancer, Parkinson's disease,
PT   asthma, or infections.
XX
PS   Claim 20; SEQ ID NO 49; 447pp; English.
XX
CC   The invention relates to a novel isolated NOVX polypeptide. The
CC   polypeptide of the invention demonstrates, antidiabetic, anorectic,
CC   cardiant, hypotensive, antiarteriosclerotic, virucide, antibacterial,
CC   fungicide, protozoacide, nootropic, neuroprotective, antiparkinsonian,
CC   anticonvulsant, osteopathic, antiarthritic, antiinflammatory,
CC   dermatological, antiasthmatic and antilipaemic activities. The
CC   polypeptides, nucleic acid molecules and antibodies may be useful for
CC   treating or diagnosing diseases including metabolic disorders such as
CC   diabetes and obesity, infectious diseases, anorexia, cancer,
CC   cardiovascular diseases including hypertension and atherosclerosis,
CC   neurodegenerative disorders such as Alzheimer's disease, Parkinson's
CC   disease and epilepsy, immune disorders e.g. osteoarthritis, haemopoietic
CC   disorders, inflammatory skin disorders, asthma and dyslipidaemia.
CC   Furthermore, the nucleic acids and polypeptides may also be used to
CC   identify molecules that modulate or inhibit neurogenesis, cell
CC   differentiation and proliferation, haemopoiesis, wound healing and
CC   angiogenesis, as well as in gene therapy. Finally, the nucleic acids may
CC   be used as hybridisation probes, in chromosome mapping, tissue typing,
CC   preventive medicine and pharmacogenomics. The current sequence is that of
CC   the human NOV cDNA of the invention.
XX
SQ   Sequence 1733 BP; 291 A; 558 C; 546 G; 338 T; 0 U; 0 Other;
```

```
Query Match        79.2%;  Score 1209;  DB 10;   Length 1733;
Best Local Similarity   89.8%;  Pred. No. 3.6e-250;
Matches 1372;  Conservative   0;  Mismatches   0;  Indels 150;  Gaps   1;

Qy      1 ATGGACCACCTCGGGGCGTCCCTCTGGCCCAGTCGGCGTCCCTTGTCTCCTGCTCGCT   60
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    145 ATGGACCACCTCGGGGCGTCCCTCTGGCCCAGTCGGCGTCCCTTGTCTCCTGCTCGCT  204

Qy     61 GGGGCGCCTGGGCGCGCCCCCGCTAACCTCCCGACCCCAAGTTCGAGAGCAAAGCGGCC  120
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    205 GGGGCGCCTGGGCGCGCCCCCGCTAACCTCCCGACCCCAAGTTCGAGAGCAAAGCGGCC  264

Qy    121 TTGCTGGGCGGGCCCGGGGCCGAAGAGCTTCTGTGCTTCACCGAGCGGTTGGAGGACTTG  180
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    265 TTGCTGGGCGGGCCCGGGGCCGAAGAGCTTCTGTGCTTCACCGAGCGGTTGGAGGACTTG  324

Qy    181 GTGTGTTTCTGGGAGGAAGCGGGCAGCGGCTGGGGTGGGCCCGGGCAACTACAGCTTCTCC  240
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    325 GTGTGTTTCTGGGAGGAAGCGGGCAGCGGCTGGGGTGGGCCCGGGCAACTACAGCTTCTCC  384

Qy    241 TACCAGCTCGAGGATGAGCCATGGAAGTGTGTCGCCTGCACCAGGCTCCCACGGCTCGT  300
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    385 TACCAGCTCGAGGATGAGCCATGGAAGTGTGTCGCCTGCACCAGGCTCCCACGGCTCGT  444

Qy    301 GGTGCGGTGGCGCTTCGTGTGTTCTCGTGTTCTGCTCGCTACAGCGACACGTCGAGCTTCGTCGTGCCCCTA  360
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    445 GGTGCGGTGGCGCTTCGTGTGTTCTCGTGTTCTGCTCGCTACAGCGACACGTCGAGCTTCGTCGTGCCCCTA  504

Qy    361 GAGTTGCGGTCACAGCAGCCTCCGGCGCTCCGGCCTCCGGCGATATCACCGTGTCATCCACATCAAT  420
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    505 GAGTTGCGGTCACAGCAGCCTCCGGCGCTCCGGCCTCCGGCGATATCACCGTGTCATCCACATCAAT  564

Qy    421 GAAGTAGTGCTCCTAGACGCGCCCCGTGGGGCGCTGGTGTGCTGACGAGAGCGGC  480
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    565 GAAGTAGTGCTCCTAGACGCGCCCCGTGGGGCGCTGGTGTGCTGACGAGAGCGGC  624

Qy    481 CACGTAGTTGTTGGCGCTGGCTCCCGGCCTGGAGACACCCATGACTCTCACATCCGCTAC  540
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    625 CACGTAGTTGTTGGCGCTGGCTCCCGGCCTGGAGACACCCATGACTCTCACATCCGCTAC  684

Qy    541 GAGGTGGACGTCTCGGCGCAACGGCAGGAGCGTACAGAGGGTGGAGATCCTGGAG  600
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    685 GAGGTGGACGTCTCGGCGCAACGGCAGGAGCGTACAGAGGGTGGAGATCCTGGAG  744

Qy    601 GGCCGCACCGAGTGTGCTGACGAACCTGCGGGGCCGGGGCCGTACACCTTCGCCGTC  660
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    745 GGCCGCACCGAGTGTGCTGACGAACCTGCGGGGCCGGGGCCGTACACCTTCGCCGTC  804

Qy    661 CGGGCGGCGTATGCGTGAGCGAGCTTCGGCGGCTTCGGCGGCCTGGAGCCGTG  720
          |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    805 CGGGCGGCGTATGCGTGAGCGAGCTTCGGCGGCTTCGGCGGCCTGGAGCCGTG  864

Qy    721 TCGCTGCTGACGCGCCTAGGGACCCGTGGACCTGACCTGGACCCCGACCCTCGTG  780
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    865 TCGCTGCTGACGCGCCTAGGGACCCGTGGACCTGACCTGGACCCCGACCCTCGTG  924

Qy    781 GTCATCCTGGTGTGCTGCACCGTGCTCCACCGGGCTCTGAAGCAG  840
          |||||||||||||||||||||||||||||||||||||||||||||
Db    925 GTCATCCTGGTGTGCTGCACCGTGCTCCACCGGGCTCTGAAGCAG  984

Qy    841 AAGGTAACTTCCAGCTGGCTGTACCAGAATGATGGGCTCCACCACCAC  900
          |||||||||||||||||||||||||||||||||||||||||||||||||
Db    985 AAGGTAACTTCCAGCTGGCTGTACCAGAATGATGGGCTCTTCACCACCCAC  1044

Qy    901 AAGGTAACTTCCAGCTGGCTGTACCAGAATGATGGGCTCTGG  960
          ||||||||||||||||||||||||||||||||||||||||||
Db   1045 AAGGTAACTTCCAGCTGGCTGTACCAGAATGATGGGCTCCTGC  1104

Qy    961 ACCCCCTTCACGGAGGAGCTTCCTCAGAGCGTCCTCTCAGAGGGG  1020
          |||||||||||||||||||||||||||||||||||||||||||||
Db   1105 ACCCCCTTCACGGAGGAGCTTCCTCAGAGCGTCCTCTCAGAGGGG  1164
```

```
Qy   1021 ACGATGCAGGCAGTGGAGCCGGGGACAGATGATGAGGGCCCCCTGCTGGAGCCAGTGGGC 1080
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1165 ACGATGCAGGCAGTGGAGCCGGGGACAGATGATGAGGGCCCCCTGCTGGAGCCAGTGGGC 1224

Qy   1081 AGTGAGCATGCCCAGGATACCTATCTGGTGCTGGACAAATGGTTGCTGCCCCGGAACCCG 1140
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1225 AGTGAGCATGCCCAGGATACCTATCTGGTGCTGGACAAATGGTTGCTGCCCCGGAACCCG 1284

Qy   1141 CCCAGTGAGGACCTCCCAGGGCCTGGTGGCAGTGTGGACATAGTGGCCATGATGAAGGC 1200
          |||||||||||||||||
Db   1285 CCCAGTGAGGACCTCCCAGGG------------------------------------- 1305

Qy   1201 TCAGAAGCATCCTCCTGCTCATCTGCTTTGGCCTCGAAGCCCAGCCCAGAGGGAGCCTCT 1260
          
Db   1306 ---------------------------------------------------------- 1305

Qy   1261 GCTGCCAGCTTTGAGTACACTATCCTGGACCCGAGCTCCCAGCTCTTGCGTCCATGGACA 1320
          |||||  ||
Db   1306 -----------------------------------------CCATGGGCA 1314

Qy   1321 CTGTGCCCTGAGCTGCCCCCTACCCCACCCTAAAGTACCTGTACCTTGTGTGGTATCT 1380
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1315 CTGTGCCCTGAGCTGCCCCCTACCCCACCCCACCCTAAAGTACCTGTACCTTGTGTGGTATCT 1374

Qy   1381 GACTCTGGCATCTCAACTGACTACAGCTCAGGGGACTCCCAGGGGAGCCCAAGGGGGCTTA 1440
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1375 GACTCTGGCATCTCAACTGACTACAGCTCAGGGGACTCCCAGGGGAGCCCAAGGGGGCTTA 1434

Qy   1441 TCCGATGGCCCCTACTCCAACCCTTATGAGAACAGCCCTTATCCCAGCCGCTGAGCCCTG 1500
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1435 TCCGATGGCCCCTACTCCAACCCTTATGAGAACAGCCCTATCCCAGCCCTGAGCCCTCTG 1494

Qy   1501 CCCCCAGCTATGTGGCTTGCTCTTAG 1527
          ||||||||||||||||||||||||||
Db   1495 CCCCCAGCTATGTGGCTTGCTCTTAG 1521
```

Search completed: April 20, 2005, 21:30:47
Job time : 856.5 secs

```
          GenCore version 5.1.6
      Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on:        April 20, 2005, 19:31:21 ; Search time 6669 Seconds
                                (without alignments)
                                11094.784 Million cell updates/sec

Title:         US-09-016-159D-4
Perfect score: 1527
Sequence:      1 atgaccaactcgggggctc.........gctatgtggctgctcttag 1527

Scoring table: IDENTITY.NUC
               Gapop 10.0 , Gapext 1.0

Searched:      4708233 seqs, 24227607955 residues

Total number of hits satisfying chosen parameters:      9416466

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

                                                SUMMARIES

Database :     GenEmbl:*
          1: gb_ba:*
          2: gb_htg:*
          3: gb_in:*
          4: gb_om:*
          5: gb_ov:*
          6: gb_pat:*
          7: gb_ph:*
          8: gb_pl:*
          9: gb_pr:*
         10: gb_ro:*
         11: gb_sts:*
         12: gb_sy:*
         13: gb_un:*
         14: gb_vi:*
```

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

```
                *
Result          Query
No.     Score   Match  Length  DB  ID          Description
-----   -----   -----  ------  --  ----------  -----------------
  1     1527    100.0   1624    9  HUMERPR     M34986 Human eryth
  2     1527    100.0   1818    6  AX658275    AX658275 Sequence
  3     1527    100.0   1818    6  HUMERYTH    M60459 Human eryth
  4    1525.4    99.9   1527    6  AR062465    AR062465 Sequence
  5    1517.4    99.4   1848    6  CQ724319    CQ724319 Sequence
  6    1312.5    85.9   1317    6  AR202361    AR202361 Sequence
  7    1312.2    85.9   1317    6  AX008198    AX008198 Sequence
  8    1307.4    85.6   1317    6  AR202362    AR202362 Sequence
  9    1307.4    85.6   1317    6  AX008201    AX008201 Sequence
 10    1276.6    83.6   2154    9  BC019092    BC019092 Homo sapi
 11    1276.6    83.6   5761    9  AK074082    AK074082 Homo sapi
 12    1155.0    75.7   1843    9  AF274305    AF274305 Sus scrof
 13    1080.6    70.8   1741   10  MUSERPR     J04843 Mouse eryth
 14    1078.6    70.6   1695   10  BC003953    BC003953 Mus muscu
 15    1075.0    70.5   1769   10  BC046282    BC046282 Mus muscu
 16    1074.2    70.3   1859   10  S59388      S59388 Mus sp. ery
 17    1058.2    69.3   1524    4  RATEPOR     D13566 Rattus norv
 18     877.2    57.4   1164    4  AY029232    AY029232 Ovis arie
 19     810.4    53.1   5683    6  AX554424    AX554424 Sequence
 20     810.4    53.1   6122    6  AX554413
 21     810.4    53.1   6122    6  AX554422
 22     810.4    53.1   6122    6  AX642149
 23     810.4    53.1   6125    6  AX554420
 24     810      53.0   4883    6  AX600116
 25    774.2     50.7   1257    6  AY029231
 26    750.4     49.1   4990    6  AR119365
 27     750      49.1   750     6  AR031378
 28     750      49.1   750     6  BD134377
 29     750      49.1   750     6  AR428937
 30     750      49.1   750     6  BD009746
 31    748.2     49.0   6256    6  BD168201
 32     741      48.5   5565    6  BD168199
 33     702      46.0   1666    6  BD090961
 34    615.4     40.3   8647    9  S45332
 35    615.4     40.3  119638    9  AC024575
 36    615.4     40.3  156897    2  AC018782
 37    615.4     40.3  217668    2  AC020561
 38     545      35.7  249366    2  AC151886
 39    524.4     34.3   686      4  BIU61398
 40    524.4     34.3   686      4  BTU61399
 41    517.2     33.9   951     10  D83509
 42    514.4     33.7   824      6  E06723
 43     478      31.3  236017    2  AC150437
 44    432.4     28.3   6561    10  MMERYPR
 45    432.4     28.3  216266    2  AC073797
```

```
                                        A
                                       AX
                                       AX6
                                       AX5
                                       AX600
                                       AY029
                                       AR119
                                       AR0313
                                       BD134377
                                       AR428937  Seq
                                       BD009746  Compositi
                                       BD168201  Method of
                                       BD168199  Method of
                                       BD090961  Assay emp
                                       S45332    erythropoie
                                       AC024575  Homo sapi
                                       AC018782  Homo sapi
                                       AC020561  Homo sapi
                                       AC151886  Saimiri s
                                       BIU61398  Bos indicus
                                       BTU61399  Bos taurus
                                       D83509    Rattus norv
                                       E06723    cDNA encodi
                                       AC150437  Lemur cat
                                       X53081    Mouse gene
                                       AC073797  Mus muscu
```

                                                ALIGNMENTS

```
RESULT 1
HUMERPR
LOCUS       HUMERPR                  1624 bp    mRNA    linear    PRI 08-NOV-1994
DEFINITION  Human erythropoietin receptor mRNA, complete cds.
ACCESSION   M34986
VERSION     M34986.1  GI:182200
KEYWORDS    erythropoietin receptor.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 1624)
  AUTHORS   Ehrenman,K. and St John,T.
  TITLE     Sequence of the human erythropoietin gene
  JOURNAL   Unpublished (1990)
  COMMENT   Original source text: Human cell line OCI-M1, cDNA to mRNA.
            Draft entry and computer-readable sequence for [1] kindly submitted
            by T.St.John, 07-JUN-1990.
            Author address: T.St John
            Hutchinson Cancer Res Center
            1124 Columbia St.
            Seattle, WA
            Email: TSTJOHN@FHCRCVAX.

FEATURES             Location/Qualifiers
     source          1..1624
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /map="19p13.2"
     gene            1..1624
                     /gene="EPOR"
     CDS             1..1527
                     /gene="EPOR"
                     /note="erythropoietin receptor precursor"
                     /codon_start=1
                     /protein_id="AAA52401.1"
                     /db_xref="GI:182201"
                     /db_xref="GDB:G00-125-242"
                     /translation="MDHLGASLWPQVGSLCLLLAGAAWAPPPNLPDPKFESKAALLAA
                     RGPEELLCFTERLEDLVCFWEEAASAGVGPGNYSFSYQLEDEPWKLCRLHQAPTARGA
                     VRFWCSLPTADTSSFVPLELRVTAASGAPRYHRVIHINEVVLLDAPVGLVARLADESG
                     HVVLRWLPPPETPMTSHIRYEVDVSAGNGAGSVQRVEILEGRTECVLSNLRGRTRYTF"
```

```
                 AVRARMAEPSFGGFWSAWSEPVSLLTPSDLDPLILTLSLILVVILVLLTVLALLSHRR
                 ALKQKIWPGISPESEFPEGLFTTHKGNFQLWLYQNDGCLWWSPCTPFTEDPPASLEVL
                 SERCWGTMQAVEPGTDDEGPLLEPVGSEHAQDTYLVLDKWLLPRNPPSEDLPGPGGSV
                 DIVAMDEGSEASSCSSALASKPSPEGASAASPEYTIIDPSSQLLRPWTLCPELPPTPP
                 HLKYLYLVVSDSGISTDYSSGDSQAQGGLSDGPYSNPYENSLIPAAEPLPPSYVACS
                 "
sig_peptide      1..72
                 /gene="EPOR"
                 /note="erythropoietin receptor signal peptide"
mat_peptide      73..1524
                 /gene="EPOR"
                 /product="erythropoietin receptor; G00-125-242"

ORIGIN

Query Match      100.0%;   Score 1527;   DB 9;   Length 1624;
Best Local Similarity   100.0%;   Pred. No. 1.5e-257;
Matches 1527;   Conservative   0;   Mismatches   0;   Indels   0;   Gaps   0;

Qy      1 ATGGACCACCTCGGGGCGTCCCCTGGCCCAGTGCGGTCGGCTCCCTTGTCTCCTGCTGCT 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 ATGGACCACCTCGGGGCGTCCCCTGGCCCAGTGCGGTCGGCTCCCTTGTCTCCTGCTGCT 60

Qy     61 GGGGCCGCCTGGGCCCCGGGGGCGCCCCGCTCAACCTCCCGGACCCCAAGTTCGAGAGCAAAGCGGCC 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61 GGGGCCGCCTGGGCCCCGGGGGCGCCCCGCTCAACCTCCCGGACCCCAAGTTCGAGAGCAAAGCGGCC 120

Qy    121 TTGCTGGCGGGCCCCGGGGGCCCGAAGAGCTTCTGTGCTTCACCGAGCCGGTTGGAGGACTTG 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121 TTGCTGGCGGGCCCCGGGGGCCCGAAGAGCTTCTGTGCTTCACCGAGCCGGTTGGAGGACTTG 180

Qy    181 GTGTGTTTCTGGAGAAGCGGCGAGCGGCTGGGGTGGGCCCGGGCAACTACAGCTTCTCC 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    181 GTGTGTTTCTGGAGAAGCGGCGAGCGGCTGGGGTGGGCCCGGGCAACTACAGCTTCTCC 240

Qy    241 TACCAGTCTGAGGATGAGCCATGGACGTGTGTGCGCTGGCCACCAGCTCCCACGGCTTCGT 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    241 TACCAGTCTGAGGATGAGCCATGGACGTGTGTGCGCTGGCCACCAGCTCCCACGGCTTCGT 300

Qy    301 GGTGCGGGTGCGCTTCGCACAGCAGCCTTCGGGGCGCTCCGGGCTCCGAGCTCGTCGTCGCCCCTA 360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    301 GGTGCGGGTGCGCTTCGCACAGCAGCCTTCGGGGCGCTCCGGGCTCCGAGCTCGTCGTCGCCCCTA 360

Qy    361 GAGTTGGCGGTCACAGCAGCCTTCGGCGGCGTCCGCGATATCACCGTGTCATCCACATCAAT 420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    361 GAGTTGGCGGTCACAGCAGCCTTCGGCGGCGTCCGCGATATCACCGTGTCATCCACATCAAT 420

Qy    421 GAAGTAGTGCTCCTAGACGCCCCGTGGGCGGTGGTGCGGTGACGAGAGCGGC 480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    421 GAAGTAGTGCTCCTAGACGCCCCGTGGGCGGTGGTGCGGTGACGAGAGCGGC 480

Qy    481 CACGTAGTGTTCGCGCTGCTGCTCCCGCCTGAGACACACCATGAGCGTCTCACATCCGCTAC 540
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    481 CACGTAGTGTTCGCGCTGCTGCTCCCGCCTGAGACACACCATGAGCGTCTCACATCCGCTAC 540

Qy    541 GAGGTGACGCGTCTCGGCCGCGCAACGGGCAGGAGCCTACAGAGAGGTGGAGATCCTGGAG 600
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    541 GAGGTGACGCGTATGCGCTCTCGGCCGCGCAACGGGCAGGAGCCTACAGAGAGGTGGAGATCCTGGAG 600

Qy    601 GGCCGCACCGAGTGTGTGCTGCAACCTGCGGGGCGCGGACGCGCTACACCTTCGCCGTC 660
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    601 GGCCGCACCGAGTGTGTGCTGCAACCTGCGGGGCGCGGACGCGCTACACCTTCGCCGTC 660

Qy    661 CCGCGGCCCGTATGCTGAGCCGAGCTTCGGCGGCTTCGGAGCGCGCCTGGTCGGAGCCGTGTG 720
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    661 CGGCGGCCCGTATGCTGAGCCGAGCTTCGGCGGCTTCGGAGCGCGCCTGGTCGGAGCCGTGTG 720

Qy    721 TCGCTGCTCGACGCGACCCTGCTGACCCTGCTGCTGCTGGAGTCTCCCTGAGCTCCCTCATCCTCGTG 780
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    721 TCGCTGCTCGACGCGACCCTGCTGACCCTGCTGCTGCTGGAGTCTCCCTGAGCTCCCTCATCCTCGTG 780

Qy    781 GTCATCCGGTCGTGCTGCTGACCGTGCTGTCGCTGCTCTCCCACCGCCGGGGCTCTGAAGCAG 840
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    781 GTCATCCGGTCGTGCTGCTGACCGTGCTGTCGCTGCTCTCCCACCGCCGGGGCTCTGAAGCAG 840
```

```
Qy    841 AAGATCTGGCCTGGCATCCCGAGCGAGAGGGCGAGTTTGAAGGCCTCTTCACCACCCAC 900
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    841 AAGATCTGGCCTGGCATCCCGAGCGAGAGGGCGAGTTTGAAGGCCTCTTCACCACCCAC 900

Qy    901 AACGGTAACTTCCAGCTGTGGCTGTACCAGAATGATGGCGTGCCTGTGTGGAGCCCTGC 960
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    901 AACGGTAACTTCCAGCTGTGGCTGTACCAGAATGATGGCGTGCCTGTGTGGAGCCCTGC 960

Qy    961 ACCCCCTCACGGAGGACCACCTGCTTCCCTGGAAGTCCTCAGAGCGCTGCTCTGTGGGG 1020
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    961 ACCCCCTCACGGAGGACCACCTGCTTCCCTGGAAGTCCTCAGAGCGCTGCTGTGGGG 1020

Qy   1021 ACGATGAGGCAGTGAGCGGGACAGATGATGAGGGCCCCTGCTGGAGCCCAGTGGGC 1080
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1021 ACGATGAGGCAGTGAGCGGGGACAGATGATGAGGGCCCCTGCTGGAGCCCAGTGGGC 1080

Qy   1081 AGTGAGCATGCCAGGATACCTATCTGTGCTGGACAAATGGTTGCTGCCCCGGAACCCG 1140
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1081 AGTGAGCATGCCAGGATACCTATCTGTGCTGGACAAATGGTTGCTGCCCCGGAACCCG 1140

Qy   1141 CCCAGTGAGGACCTCCCAGGGCCTGGCGAGTGTGGACATAGTGGCCATGGATGAAGGC 1200
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1141 CCCAGTGAGGACCTCCCAGGGCCTGGTGGCAGTGTGGACATAGTGGCCATGGATGAAGGC 1200

Qy   1201 TCAGAAGCATCCTCCTGCTCATCTGCTTGGCCTCGAAGCCCAGAGCCAGGGGAGCCTCT 1260
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1201 TCAGAAGCATCCTCCTGCTCATCTGCTTGGCCTCGAAGCCCAGAGCCAGGGGAGCCTCT 1260

Qy   1261 GCTGCCAGTTTGAGTACACTATCCTGGACCCCCACCTAAAGTACCTGTACCTTGTGGTATCT 1320
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1261 GCTGCCAGTTTGAGTACACTATCCTGGACCCCCACCTAAAGTACCTGTACCTTGTGGTATCT 1320

Qy   1321 CTGTGCCCTGAGCTGCCCCTACCCCACCCACCACAGCTCAGGGGACTCAGGGGCTTA 1380
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1321 CTGTGCCCTGAGCTGCCCCTACCCCACCCACCACAGCTCAGGGGACTCAGGGGCTTA 1380

Qy   1381 GACTCTGGCATCTCAACTGACTACAGCTTCAGGGACTCCCAGGGGAGCCCAAGGGGCTTA 1440
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1381 GACTCTGGCATCTCAACTGACTACAGCTTCAGGGACTCCCAGGGGAGCCCAAGGGGCTTA 1440

Qy   1441 TCCGATGGCCCCTACTCCAACCCTTATGAGAACAGCCTTATCCGAGAGAACAGCCTTATCCCCAGCCGCTCCTG 1500
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1441 TCCGATGGCCCCTACTCCAACCCTTATGAGAACAGCCTTATCCGAGAGAACAGCCTTATCCCCAGCCGCTCCTG 1500

Qy   1501 CCCCCAGCTATGTGGCTTGCTCTTAG 1527
          ||||||||||||||||||||||||||
Db   1501 CCCCCAGCTATGTGGCTTGCTCTTAG 1527
```

```
RESULT 2
AX658275
LOCUS       AX658275                1818 bp    DNA        linear    PAT 22-MAR-2003
DEFINITION  Sequence 191 from Patent WO03000928.
ACCESSION   AX658275
VERSION     AX658275.1  GI:29160793
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
  AUTHORS   Poulsen,H.S., Pedersen,N., Mortensen,S., Sorensen,S.B.,
            Petersen,M.W. and Elsner,H.I.
  TITLE     Methods for identification of cancer cell surface molecules and
            cancer specific promoters, and therapeutic uses thereof
  JOURNAL   Patent: WO 03000928-A 191 03-JAN-2003;
            Odin Medical A/S (DK)
FEATURES             Location/Qualifiers
     source          1..1818
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
```

ORIGIN

```
Query Match       100.0%;    Score 1527;   DB 6;   Length 1818;
Best Local Similarity  100.0%;   Pred. No. 1.5e-257;
Matches 1527;   Conservative   0;   Mismatches   0;   Indels   0;   Gaps   0;

Qy     1 ATGGACCACCTCGGGGGCTCCTCTGCGCCCCAGGTCGGCGTCCCCTTTGTCTCCTGCTCGCT   60
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   106 ATGGACCACCTCGGGGGCTCCTCTGCGCCCCAGGTCGGCGTCCCCTTTGTCTCCTGCTCGCT  165

Qy    61 GGGGCCGCCTGGGCGCCCGGGGGCGCCCCGCCCAACTCCCGGAGCCCAGTTGAGAGGCAAAGCGGGCC  120
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   166 GGGGCCGCCTGGGCGCCCGGGGGCGCCCCGCCCAACTCCCGGAGCCCAGTTGAGAGGCAAAGCGGGCC  225

Qy   121 TTGCTGGCGGCCCGGGGCCCAGAAGAGCTTCTGCTTCACCGAGCGGTTGGAGGACTTG  180
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   226 TTGCTGGCGGCCCGGGGCCCAGAAGAGCTTCTGCTTCACCGAGCGGTTGGAGGACTTG  285

Qy   181 GTGTGTTTCTGGGAGGAAGCGGCGAGCGCTGGGGTGGGCCCCGGGGCAACTCACAGCTTCTCC  240
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   286 GTGTGTTTCTGGGAGGAAGCGGCGAGCGCTGGGGTGGGCCCCGGGGCAACTCACAGCTTCTCC  345

Qy   241 TACCAGCTCGAGGATGAGAGCCATGAGAAGCTGTGTCGCGCCTACCAGGCTCCCACCGGCTCGT  300
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   346 TACCAGCTCGAGGATGAGAGCCATGAGAAGCTGTGTCGCGCCTACCAGGCTCCCACCGGCTCGT  405

Qy   301 GGTGCGGTGCGCTTCGTGTTCGCTGCCTACAGCCGACACGTCGAGCTTCGTGCGCCCTA  360
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   406 GGTGCGGTGCGCTTCGTGTTCGCTGCCTACAGCCGACACGTCGAGCTTCGTGCGCCCTA  465

Qy   361 GAGTTGCGCGTCACAGCAGCTCCGGCGCTCCGGCGATATCACCGTGTCATCCACATCAAT  420
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   466 GAGTTGCGCGTCACAGCAGCTCCGGCGCTCCGGCGATATCACCGTGTCATCCACATCAAT  525

Qy   421 GAAGTAGTGCTTCCTAGACGCGCCCGTGGGGCGGTTGGCTGACGAGGAGCGGC  480
         ||||||||||||||||||||||||||||||||||||||||||||||||||
Db   526 GAAGTAGTGCTTCCTAGACGCGCCCGTGGGGCGGTTGGCTGACGAGGAGCGGC  585

Qy   481 CACGTAGTGTTGCGCTGGCTCCCGCCCTGAGACGTCTCACATCCGCTAC  540
         ||||||||||||||||||||||||||||||||||||||||||||||||
Db   586 CACGTAGTGTTGCGCTGGCTCCCGCCCTGAGACGTCTCACATCCGCTAC  645

Qy   541 GAGGTGGACGTCTCGGCCGCCAACGGCCGGAGCCGTACAGAGAGGTGAGATCCTGGAG  600
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   646 GAGGTGGACGTCTCGGCCGCCAACGGCCGGAGCCGTACAGAGAGGTGAGATCCTGGAG  705

Qy   601 GGCCGGCACCGAGTGTGTGCTGAGCAACCTGCGGGGCGCCGGAGCCGCCTTCGCCGTC  660
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   706 GGCCGGCACCGAGTGTGTGCTGAGCAACCTGCGGGGCGCCGGAGCCGCCTTCGCCGTC  765

Qy   661 CGGCGCGTATGGCTGACCGGAGCTTCGGCCGGCTTCTGGAGCCGTGTG  720
         ||||||||||||||||||||||||||||||||||||||||||||||
Db   766 CGGCGCGTATGGCTGACCGGAGCTTCGGCCGGCTTCTGGAGCCGTGTG  825

Qy   721 TCGCTGCTGACGCGCTAGCGACCTGGACCCCCTCATCCTGACCCCCTCATCCTCGTG  780
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   826 TCGCTGCTGACGCGCTAGCGACCTGGACCCCCTCATCCTGACCCCCTCATCCTCGTG  885

Qy   781 GTCATCCTGGTGCTGCTAGCCGTGCTCCGGCTCTCCACCGCCGGGCTCTGAAGCAG  840
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   886 GTCATCCTGGTGCTGCTAGCCGTGCTCCGGCTCTCCACCGCCGGGCTCTGAAGCAG  945

Qy   841 AAGATCTGGCCTGGCATCCCGAGCCCCAGAGCGAGTTTGAAGGCTTCTTCACCACCCAC  900
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   946 AAGATCTGGCCTGGCATCCCGAGCCCCAGAGCGAGTTTGAAGGCTTCTTCACCACCCAC  1005

Qy   901 AAGGGTAACTTCCAGCTGTGGCTGTACCAGAATGATGAGCTGCCTGGTGGAGCCCCCTGC  960
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1006 AAGGGTAACTTCCAGCTGTGGCTGTACCAGAATGATGAGCTGCCTGGTGGAGCCCCCTGC  1065

Qy   961 ACCCCCTTCACGGAGACCCACCTGCTTCCCTGGAGTCCTCTCAGAGCGCTGTGGGGG  1020
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1066 ACCCCCTTCACGGAGACCCACCTGCTTCCCTGGAGTCCTCTCAGAGCGCTGTGGGGG  1125

Qy  1021 ACGATGCAGGCAGTGGAGCCGGGGACAGATGATGAGGCCCCCTGCTGCGAGCCAGTGGGC  1080
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1126 ACGATGCAGGCAGTGGAGCCGGGGACAGATGATGAGGCCCCCTGCTGCGAGCCAGTGGGC  1185

Qy  1081 AGTGAGCATGCCAGGATACCTATCTGGTGCTGGACAAATGGTGCTGCCCCGGAACCCG  1140
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1186 AGTGAGCATGCCAGGATACCTATCTGGTGCTGGACAAATGGTGCTGCCCCGGAACCCG  1245

Qy  1141 CCCAGTGAGGAGCTTCCCAGGGCCTGGTGTGGACAGTGTGGACATAGTGGCCATGATGAGGGC  1200
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1246 CCCAGTGAGGAGCTTCCCAGGGCCTGGTGTGGACAGTGTGGACATAGTGGCCATGATGAGGGC  1305

Qy  1201 TCAGAAGCATCCTCCTGCTCATCTGCTTGGCCTCGAAGCCCAGCCCAGCCTCTTCGTCATGACA  1260
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1306 TCAGAAGCATCCTCCTGCTCATCTGCTTGGCCTCGAAGCCCAGCCCAGCCTCTTCGTCATGACA  1365

Qy  1261 GCTGCCAGCTTTGAGTACATATCCTGGACCCCACCCACCCCCACCCCACCTAAAGTACCTGTGGTATCT  1320
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1366 GCTGCCAGCTTTGAGTACATATCCTGGACCCCACCCACCCCCACCCCACCTAAAGTACCTGTGGTATCT  1425

Qy  1321 CTGTGCCCTGAGCTGCCCCCTACCCCACCCACCCACCCCCACCCACTAAAGTACCTGTTGGTATCT  1380
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1426 CTGTGCCCTGAGCTGCCCCCTACCCCACCCACCCACCCCCACCCACTAAAGTACCTGTTGGTATCT  1485

Qy  1381 GACTCTGGCATCTCAACTGACTACAGCTCCAGGGGACTCCCAGGGAGCCCAAGGGGGCTTA  1440
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1486 GACTCTGGCATCTCAACTGACTACAGCTCCAGGGGACTCCCAGGGAGCCCAAGGGGGCTTA  1545

Qy  1441 TCCGATGGCCCCTACTCCAACCCTTATGAGAACAGCCTTATCCCAGCCGCTGAGCCTCTG  1500
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1546 TCCGATGGCCCCTACTCCAACCCTTATGAGAACAGCCTTATCCCAGCCGCTGAGCCTCTG  1605

Qy  1501 CCCCCCAGCTATGTGGCTTGCTCTTAG  1527
         |||||||||||||||||||||||||||
Db  1606 CCCCCCAGCTATGTGGCTTGCTCTTAG  1632
```

RESULT 3
HUMERYTH

```
LOCUS       HUMERYTH                1818 bp    mRNA    linear   PRI 08-NOV-1994
DEFINITION  Human erythropoietin receptor mRNA, complete cds.
ACCESSION   M60459
VERSION     M60459.1  GI:182244
KEYWORDS    erythropoietin receptor.
SOURCE      Homo sapiens (human)
ORGANISM    Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 1818)
AUTHORS     Jones,S.S., D'Andrea,A.D., Haines,L.L. and Wong,G.G.
TITLE       Human erythropoietin receptor: cloning, expression, and biologic
            characterization
JOURNAL     Blood 76 (1), 31-35 (1990)
MEDLINE     90304340
PUBMED      2163696
COMMENT     Original source text: Human cDNA to mRNA.
FEATURES             Location/Qualifiers
     source          1..1818
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /map="19p13.2"
     gene            1..1818
                     /gene="EPOR"
     CDS             106..1632
                     /gene="EPOR"
                     /codon_start=1
                     /product="erythropoietin receptor"
                     /protein_id="AAA52403.1"
                     /db_xref="GI:182245"
                     /db_xref="GDB:G00-125-242"
                     /translation="MDHLGASLWPQVGSLCLLLAGAAWAPPPNLPDPKFESKAALLAA
```

RGPEELLCFTERLEDLVCFWEEAAASAGVGPGNYSFSYQLEDEPWKLCRLHQAPTARGA
VRFWCSLPTADTTSSFVPLELRVTAASGAPRYHRVIHNEVVLLDAPVGLVARLADESG
HVVLRWLPPPEPTMTSHIRYEVDVSAGNGAGSVQRVEILEGRTRYTF
AVRARMAEPSFGGFWSAWSEPVSLLTPSDLDPLIILTLSLILVVILVLLTVLALLSHRR
ALKQKIWPGIPSPESEFEGLFTHKGNFQLWLYQNDGCLWWSPCTPFTEDPPASLEVL
SERCWGTMQAVEPGTDDEGPLLEPVGSEHAQDTYLVLDKWLLPRNPPSEDLPGPGGSV
DIVAMDEGSEASSCSSALASKPSPEGASAASFEYTILDPSSQLLRPWTLCPELPPTPP
HLKYLYLVVSDSGISTDYSSGDSQGAQGGLSDGPYSNPYENSLIPAAEPLPPSYVACS
"

sig_peptide   106..177
               /gene="EPOR"
               /note="G00-125-242"

mat_peptide   178..1629
               /gene="EPOR"
               /product="erythropoietin receptor"
               /note="G00-125-242"

ORIGIN

Query Match   100.0%;   Score 1527;   DB 9;   Length 1818;
Best Local Similarity 100.0%;   Pred. No.1.5e-257;
Matches 1527; Conservative   0;   Mismatches   0;   Indels   0;   Gaps   0;

Qy   1 ATGGACCACCTCGGGGCGTCCCTCTGGCCCCAGTCGGGTCTCCTTTGTCTCCTGCTGCT   60
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   106 ATGGACCACCTCGGGGCGTCCCTCTGGCCCCAGTCGGGTCTCCTTTGTCTCCTGCTGCT 165

Qy   61 GGGGCCGCCTCGGGGCCCCCGCCTAACCTCCCGACCCCAAGTTCGAGAGCAAAGCGGCC 120
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   166 GGGGCCGCCTCGGGGCCCCCGCCTAACCTCCCGACCCCAAGTTCGAGAGCAAAGCGGCC 225

Qy   121 TTGCTGGGGGCCGGGGGCCCGAAGAGTTCTGTGCTTCACCGAGCCGGTTGGAGGACTTG 180
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   226 TTGCTGGGGGCCGGGGGCCCGAAGAGTTCTGTGCTTCACCGAGCCGGTTGGAGGACTTG 285

Qy   181 GTGTGTTTTCTGGGAGGAAGCGGCGAGCGGCTGGGGTGGGCCAACTACAGCTTCTCC 240
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   286 GTGTGTTTTCTGGGAGGAAGCGGCGAGCGGCTGGGGTGGGCCAACTACAGCTTCTCC 345

Qy   241 TACCAGCTCGAGGATGAGCCATGAAGTCTGTGCCGCCTACAGCCGACACGTCCACGGCTCGT 300
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   346 TACCAGCTCGAGGATGAGCCATGAAGTCTGTGCCGCCTACAGCCGACACGTCCACGGCTCGT 405

Qy   301 GGTGCGGTGCGCTTCTGGTTCTGCTGCTTGCTGCTCCGGCCGACACGTCGAGCTTCGTGCCCCTA 360
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   406 GGTGCGGTGCGCTTCTGGTTCTGCTGCTTGCTGCTCCGGCCGACACGTCGAGCTTCGTGCCCCTA 465

Qy   361 GAGTTGCGGTCACAGCAGCCTCCGGCCGTCCGGCGTTGGGCGGTTGCGGTGTCATCCACATCAAT 420
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   466 GAGTTGCGGTCACAGCAGCCTCCGGCCGTCCGGCGTTGGGCGGTTGCGGTGTCATCCACATCAAT 525

Qy   421 GAAGTAGTGCTCCTAGACGCCCCCGTGGGGCTGTGCGCGGTTGGGCGGTGCTGACGAGAGCGGC 480
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   526 GAAGTAGTGCTCCTAGACGCCCCCGTGGGGCTGTGCGCGGTTGGGCGGTGCTGACGAGAGCGGC 585

Qy   481 CACGTAGTTGTTGCGGCTGGCCTCCGCGCCTGCACACCCATGAGACACGTCTCACATCCGCTAC 540
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   586 CACGTAGTTGTTGCGGCTGGCCTCCGCGCCTGCACACCCATGAGACACGTCTCACATCCGCTAC 645

Qy   541 GAGGTGGACGTCTCGGCCACGCCAGGGCGGCTTCGGCGGTTGGAGATCCTGGAG 600
           |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   646 GAGGTGGACGTCTCGGCCACGCCAGGGCGGCTTCGGCGGTTGGAGATCCTGGAG 705

Qy   601 GGCCGCACCGAGTTGTGTCGACGCCTTGCGGGGCGGCCGGACCGCGCTACACCTTCGCGTC 660
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   706 GGCCGCACCGAGTTGTGTCGACGCCTTGCGGGGCGGCCGGACCGCGCTACACCTTCGCGTC 765

Qy   661 CGCGCGGCTGAGCCAGCTTCGGCGGCTTCGGCGGCGGCTTCGGCGGTTCGGAGCCGTGTG 720
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   766 CGCGCGGCTGAGCCAGCTTCGGCGGCTTCGGCGGCGGCTTCGGCGGTTCGGAGCCGTGTG 825

Qy   721 TCGCTGCTGACGCCTAGGCACCTGGACCCCCTCATCCTGTGACGCTTCCCTCATCCTGTG 780
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   826 TCGCTGCTGACGCCTAGGCACCTGGACCCCCTCATCCTGTGACGCTTCCCTCATCCTGTG 885

Qy   781 GTCATCCTGGTGCTGCTGACGGTGCTCTGCGCTGCTGTGCTCCTGCTGCTGCTGGGTCTCCTGCTGGAGCTCTGAAGCAG 840
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   886 GTCATCCTGGTGCTGCTGACGGTGCTCTGCGCTGCTGTGCTCCTGCTGCTGCTGGGTCTCCTGCTGGAGCTCTGAAGCAG 945

Qy   841 AAGATCTGGCCTGGCATCCCGAGCCGCAGAGAGCGAGTTTGAAGGCCTCTCCACCACCCAC 900
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   946 AAGATCTGGCCTGGCATCCCGAGCCGCAGAGAGCGAGTTTGAAGGCCTCTTCCACCACCCAC 1005

Qy   901 AAGGGTAACTTCCAGCTGTGGCTGTACCAGAATGATGGCGTGCCTCGTGGTGGGAGCCCTGC 960
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1006 AAGGGTAACTTCCAGCTGTGGCTGTACCAGAATGATGGCGTGCCTCGTGGTGGGAGCCCCTGC 1065

Qy   961 ACCCCTTCACGGAGGACCACCTGCTTCCCTGGAGTCCTCTCAGAGTCCTCTCAGAGTGCGGGGG 1020
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1066 ACCCCCTTCACGGAGGACCACCTGCTTCCCTGGAGTCCTCTCAGAGTCCTCTCAGAGTGCTGCTGGGGG 1125

Qy   1021 ACGATGCAGGCAGTGGAGCGGGGACACGATGATGAGGGCCCCCTGCTGAGCCAGTGGGC 1080
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1126 ACGATGCAGGCAGTGGAGCGGGGACACGATGATGAGGGCCCCCTGCTGAGCCAGTGGGC 1185

Qy   1081 AGTGAGCATGCCCAGGATACCTATCTGTGCTGACAAATGGTTGCTGCCCCGGAACCCG 1140
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1186 AGTGAGCATGCCCAGGATACCTATCTGTGCTGACAAATGGTTGCTGCCCCGGAACCCG 1245

Qy   1141 CCCAGTGAGGACTCCCAGGGCCTGGTGCAGTGTGGACATAGTGGCCATGAAGGC 1200
           |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1246 CCCAGTGAGGACTCCCAGGGCCTGGTGCAGTGTGGACATAGTGGCCATGAAGGC 1305

Qy   1201 TCAGAAGCATCTCCTGTCTGCTTGGGCCTCGAAGCCCAGAGAGGAGCCTCT 1260
           ||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1306 TCAGAAGCATCTCCTGTCTGCTTGGGCCTCGAAGCCCAGAGAGGAGCCTCT 1365

Qy   1261 GCTGCCAGCTTTGAGTACACTATCCTGACCCCCAGCTCTTGCGTCCATGGACA 1320
           |||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1366 GCTGCCAGCTTTGAGTACACTATCCTGACCCCCAGCTCTTGCGTCCATGGACA 1425

Qy   1321 CTGTGCCTGAGCTGCCCCCTACCCCACCCCACCTAAAGTACCTGTGGTATCT 1380
           ||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1426 CTGTGCCTGAGCTGCCCCCTACCCCACCCCACCTAAAGTACCTGTGGTATCT 1485

Qy   1381 GACTCTGCATCTCAACTGACACAGCTCAGGGAGCCCAAGGGGCTTA 1440
           |||||||||||||||||||||||||||||||||||||||||||||
Db   1486 GACTCTGCATCTCAACTGACACAGCTCAGGGAGCCCAAGGGGCTTA 1545

Qy   1441 TCCGATGGCCCCTACTCCAACCCTTATGAGAACAGCCTTATCCCAGCCTGAGCCTCTG 1500
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1546 TCCGATGGCCCCTACTCCAACCCTTATGAGAACAGCCTTATCCCAGCCTGAGCCTCTG 1605

Qy   1501 CCCCCAGCTATGTGGCTTGCTCTTAG 1527
           ||||||||||||||||||||||||||
Db   1606 CCCCCAGCTATGTGGCTTGCTCTTAG 1632

RESULT 4
AR062465
LOCUS   AR062465   1527 bp   DNA   linear   PAT 29-SEP-1999
DEFINITION   Sequence 4 from patent US 5843726.
ACCESSION   AR062465
VERSION   AR062465.1   GI:5990156
KEYWORDS
SOURCE   Unknown.
  ORGANISM   Unknown.
              Unclassified.
REFERENCE   1 (bases 1 to 1527)
  AUTHORS   Lee,J.Y.
  TITLE   Expression of a cleavable fusion protein comprising a soluble human
              erythropoietin receptor protein fragment
  JOURNAL   Patent: US 5843726-A 4 01-DEC-1998;
FEATURES   Location/Qualifiers
     source   1..1527
                     /organism="unknown"
                     /mol_type="unassigned DNA"

ORIGIN

```
Query Match        99.9%;   Score 1525.4;  DB 6;  Length 1527;
Best Local Similarity  99.9%;   Pred. No. 2.9e-257;
Matches 1526;   Conservative   0;  Mismatches   1;  Indels   0;  Gaps   0;

Qy      1 ATGGACCACCTCGGGGCGGTCCCTGCCCCAGGTCGGCTCCCTTTGTCTCCTGCTCGCT 60
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 ATGGACCACCTCGGGGCGGTCCCTGCCCCAGGTCGGCTCCCTTTGTCTCCTGCTCGCT 60

Qy     61 GGGGCCGGCTGGGCGCCCCCGCTAACCTCCCGGACCCCAAGTTCGAGAGACAAAGCGGCC 120
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61 GGGGCCGGCTGGGCGCCCCCGCTAACCTCCCGGACCCCAAGTTCGAGAGACAAAGCGGCC 120

Qy    121 TTGCTGGGCGCCCGGGGGCCCGAAGAGTTCTGTGCTTCACCGAGCCGGTTGGAGGACTTG 180
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121 TTGCTGGGCGCCCGGGGGCCCGAAGAGTTCTGTGCTTCACCGAGCCGGTTGGAGGACTTG 180

Qy    181 GTGTGTTTCTGGGAGGAAGCGGCGAGCGTGGGGTGGGCCCGGGCAACTACAGCTTCTCC 240
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    181 GTGTGTTTCTGGGAGGAAGCGGCGAGCGTGGGGTGGGCCCGGGCAACTACAGCTTCTCC 240

Qy    241 TACCAGTCGAGGATGAGCCATGGAAGCTGTGTCGCCTACAGCCGCACCGGTCTCCACGGCTCGT 300
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    241 TACCAGTCGAGGATGAGCCATGGAAGCTGTGTCGCCTACAGCCGCACCGGTCTCCACGGCTCGT 300

Qy    301 GGTCGGCGGTGCGCTTCTGTGTTCGCGGCGCTTCGCTGCCTCGAGCTTCGTGCCCCTA 360
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    301 GGTCGGCGGTGCGCTTCTGTGTTCGCGGCGCTTCGCTGCCTCGAGCTTCGTGCCCCTA 360

Qy    361 GAGTTGCGGTCACAGCAGCCTCCGGGCGTCCGCGATATCACCGTGTCATCCACATCAAT 420
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    361 GAGTTGCGGTCACAGCAGCCTCCGGGCGTCCGCGATATCACCGTGTCATCCACATCAAT 420

Qy    421 GAAGTAGTGCTCCTAGACGCCCCCGTGGGGCGTTGGTGGGCGGTTGGCTGACGAGAGCGGC 480
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    421 GAAGTAGTGCTCCTAGACGCCCCCGTGGGGCGTTGGTGGGCGGTTGGCTGACGAGAGCGGC 480

Qy    481 CACGTAGTGTTGCGCTGGCTGCTCCCGCGCCGCCTGAGACGTCCATCCGCTAC 540
          |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    481 CACGTAGTGTTGCGCTGGCTGCTCCCGCGCCGCCTGAGACACCCATCCGCTAC 540

Qy    541 GAGGTGGACGTCTCGGCGCCGGCAACGGCGGAGCGTACAGAGAGGTGGAGATCCTGGAG 600
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    541 GAGGTGGACGTCTCGGCGCCGGCAACGGCGGAGCGTACAGAGAGGTGGAGATCCTGGAG 600

Qy    601 GGCCGGCACCGAGTGTGTGCTGAGCAACCTGCGGGCGCCGGACGCGCTACACCTTCGCGTC 660
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    601 GGCCGGCACCGAGTGTGTGCTGAGCAACCTGCGGGCGCCGGACGCGCTACACCTTCGCGTC 660

Qy    661 CGGCCGCGGTATGGCTGAGCCGGCAGCTTCGGAGCCGTCGGAGCCGTGTG 720
          |||||||||||||||||||||||||||||||||||||||||||||||||||
Db    661 CTGCCGCGGTATGGCTGAGCCGGCAGCTTCGGCCGGTCGGAGCCGTGTG 720

Qy    721 TCGCTGCTGACGCGCCACCTGGACCTGGACGCGCTCTCCATCCTCCGTG 780
          |||||||||||||||||||||||||||||||||||||||||||||||||||
Db    721 TCGCTGCTGACGCGCCACCTGGACCTGGACGCGCTCTCCATCCTCCGTG 780

Qy    781 GTCATCCTGGTGCTGTGACCGTGCTGCTCCCACCGCCGGGCTCTGAAGCAG 840
          |||||||||||||||||||||||||||||||||||||||||||||||||||
Db    781 GTCATCCTGGTGCTGTGACCGTGCTGCTCCCACCGCCGGGCTCTGAAGCAG 840

Qy    841 AAGAGTCTGGCTGGCATCCCAGAGCGCCAGAGTTGAAGGCCTCTTCACCACCCAC 900
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    841 AAGAGTCTGGCCTGGCATCCCAGAGCCCAGAGTTGAAGGCCTCTTCACCACCCAC 900

Qy    901 AAGGGTAACTTCAGCTGTGTGGCTGTACCAGAATGATGGCTGCGTGTGGTGGAGCCCCTGC 960
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    901 AAGGGTAACTTCAGCTGTGTGGCTGTACCAGAATGATGGCTGCGTGTGGTGGAGCCCCTGC 960

Qy    961 ACCCCCTTCACGGAGGACCCACCTGCTTCCCTGGAAGTCTTCCTGGAAGTCTTCAGAGCGCGTGCTGTGGGG 1020
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    961 ACCCCCTTCACGGAGGACCCACCTGCTTCCCTGGAAGTCTTCCTGGAAGTCTTCAGAGCGCGTGCTGTGGGG 1020
```

```
Qy   1021 ACGATGCAGGCAGTGGAGCCGGGGACAGATGATGAGGGCCCCCTGCTGGAGCCAGTGGGC 1080
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1021 ACGATGCAGGCAGTGGAGCCGGGGACAGATGATGAGGGCCCCCTGCTGGAGCCAGTGGGC 1080

Qy   1081 AGTGAGCAGTGCATGCCCAGGATACCATCTGTGCTGTGACAAATGGTGTCTGCCCGAACCCG 1140
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1081 AGTGAGCAGTGCATGCCCAGGATACCATCTGTGCTGTGACAAATGGTGTCTGCCCGAACCCG 1140

Qy   1141 CCCAGTGAGGACTTCCCAGGCCTGTGGCAGTGTGACATAGTGGCCATGATGAAGGC 1200
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1141 CCCAGTGAGGACTTCCCAGGCCTGTGGCAGTGTGACATAGTGGCCATGATGAAGGC 1200

Qy   1201 TCAGAAGCATCCTCCTGCTCATCTGCTTGGCCTCGAAGCCCAGCCCAGAGGGAGCCCTCT 1260
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1201 TCAGAAGCATCCTCCTGCTCATCTGCTTGGCCTCGAAGCCCAGCCCAGAGGGAGCCCTCT 1260

Qy   1261 GCTGCCAGCTTTGAGTACACATATCCTGGACCCCCAGCTCTTGCGTCCATGGACA 1320
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1261 GCTGCCAGCTTTGAGTACACATATCCTGGACCCCCAGCTCTTGCGTCCATGGACA 1320

Qy   1321 CTGTGCCCTGAGCTGCCCCTACCCCACCTAAAGTACCTGTACCTTGTGGTATCT 1380
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1321 CTGTGCCCTGAGCTGCCCCTACCCCACCTAAAGTACCTGTACCTTGTGGTATCT 1380

Qy   1381 GACTCTGCATCTCAACTGACTACAGCTTCAGGGGACTCAGGGAGCCCAAGGGGCTTA 1440
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1381 GACTCTGCATCTCAACTGACTACAGCTTCAGGGGACTCAGGGAGCCCAAGGGGCTTA 1440

Qy   1441 TCCGATGGCCCCTACTCCAACCCTTATGAGAACAGCAGCCTTATCCCAGCCGCTGAGCCTCTG 1500
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1441 TCCGATGGCCCCTACTCCAACCCTTATGAGAACAGCAGCCTTATCCCAGCCGCTGAGCCTCTG 1500

Qy   1501 CCCCCAGCTATGTGGCTTGCTCTTAG 1527
          ||||||||||||||||||||||||||
Db   1501 CCCCCAGCTATGTGGCTTGCTCTTAG 1527
```

RESULT 5
```
LOCUS        CQ724319                    1848 bp     DNA           linear    PAT 03-FEB-2004
DEFINITION   Sequence 10253 from Patent WO02068579.
ACCESSION    CQ724319
VERSION      CQ724319.1   GI:42285176
KEYWORDS     .
SOURCE       Homo sapiens (human)
  ORGANISM   Homo sapiens
             Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
             Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE    1  (bases 1 to 1848)
  AUTHORS    Venter,C.J., Adams,M.C., Li,P.W. and Myers,E.W.
  TITLE      Kits, such as nucleic acid arrays, comprising a majority of
             humanexons or transcripts, for detecting expression and other uses
             thereof
  JOURNAL    Patent: WO 02068579-A 10253  06-SEP-2002;
             PE Corporation (NY) (US)
FEATURES             Location/Qualifiers
     source          1..1848
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
ORIGIN
```

```
Query Match        99.4%;   Score 1517.4;  DB 6;  Length 1848;
Best Local Similarity  99.6%;   Pred. No. 6.9e-256;
Matches 1521;   Conservative   0;  Mismatches   6;  Indels   0;  Gaps   0;

Qy      1 ATGGACCACCTCGGGGCGGTCCCTGCCCCAGGTCGGCTCCCTTTGTCTCCTGCTCGCT 60
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    137 ATGGACCACCTCGGGGCGGTCCCTGCCCCAGGTCGGCTCCCTTTGTATACCTGCTCGCT 196

Qy     61 GGGGCCGGCTGGGCGCCCCCGCTAACCTCCCGGACCCCAAGTTCGAGAGACAAAGCGGCC 120
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61 GGGGCCGGCTGGGCGCCCCCGCTAACCTCCCGGACCCCAAGTTCGAGAGACAAAGCGGCC 120
```

```
RESULT 6
AR202361
LOCUS       AR202361              1317 bp    DNA        linear   PAT 20-APR-2002
DEFINITION  Sequence 4 from patent US 6361998.
ACCESSION   AR202361
VERSION     AR202361.1  GI:20256900
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown
            Unclassified.
REFERENCE   1  (bases 1 to 1317)
  AUTHORS   Bell,D.N., Matthews,K.Emma. and Mueller,S.G.
  TITLE     Efficient culture of stem cells for the production of hemoglobin
  JOURNAL   Patent: US 6361998-A 4 26-MAR-2002;
FEATURES             Location/Qualifiers
     source          1..1317
                     /organism="unknown"
                     /mol_type="unassigned DNA"
ORIGIN

Query Match             85.9%;  Score 1312.2;  DB 6;  Length 1317;
Best Local Similarity   99.8%;  Pred. No. 6.3e-220;
Matches 1314;  Conservative 0;  Mismatches 3;  Indels 0;  Gaps 0;
```

```
SOURCE      Homo sapiens (human)
ORGANISM    Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
AUTHORS     Matthews,K.E., Bell,D. and Mueller,S.G.
TITLE       The efficient culture of stem cells for the production of
            hemoglobin
JOURNAL     Patent: WO 9967360-A 4 29-DEC-1999;
            HEMOSOL INC (CA)
FEATURES    Location/Qualifiers
  source    1..1317
            /organism="Homo sapiens"
            /mol_type="unassigned DNA"
            /db_xref="taxon:9606"
  terminator 1315..1317

ORIGIN
Query Match          85.9%;   Score 1312.2;   DB 6;   Length 1317;
Best Local Similarity 99.8%;   Pred. No. 6.3e-220;
Matches 1314;   Conservative    0;   Mismatches    3;   Indels    0;   Gaps    0;

Qy       1 ATGGACCACCTCGGGGCGTCCCTCTGGCCCCAGTCGGCTCCCTTGTCTCCTGCTCGCT    60
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1 ATGGACCACCTCGGGGCGTCCCTCTGGCCCCAGTCGGCTCCCTTGTCTCCTGCTCGCT    60

Qy      61 GGGGCCGCCTGGGCGCCCGGGGGCCCGAAGAGTTCGAGAGCAAAGCGGGC   120
           ||||||||||||||||||||||||||||||||||||||||||||||||||
Db      61 GGGGCCGCCTGGGCGCCCGGGGGCCCGAAGAGTTCGAGAGCAAAGCGGGC   120

Qy     121 TTGCTGCGGCGCCCGGGGCCCGAAGAGCTTCTGTGCTTCACCGAGCGGTTGAGGACTTG   180
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     121 TTGCTGCGGCGCCCGGGGCCCGAAGAGCTTCTGTGCTTCACCGAGCGGTTGAGGACTTG   180

Qy     181 GTGTGTTTCTGGGAGGAAGCGGCGAGCGCTGGGGGCCCGGCCAACTACAGCTTCTCC   240
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     181 GTGTGTTTCTGGGAGGAAGCGGCGAGCGCTGGGGGCCCGGCCAACTACAGCTTCTCC   240

Qy     241 TACCAGCTCGAGGATGAGCCAGCGCCTCCTGCTGCCTACAGCCGACACGTCGTGTCGT   300
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     241 TACCAGCTCGAGGATGAGCCAGCGCCTCCTGCTGCCTACAGCCGACACGTCGTGTCGT   300

Qy     301 GGTGCGGTGCCTTCTGGTGTTCGCTGCCTACAGCCGGATATCACCGGCTCGT   360
           |||||||||||||||||||||||||||||||||||||||||||||||||||
Db     301 GGTGCGGTGCCTTCTGGTGTTCGCTGCCTACAGCCGGATATCACCGGCTCGT   360

Qy     361 GAGTTGCGCGTCTCACAGAGCAGCCTCCGGCGGTTGTGGCGCGGTGACGAGAGCGGC   420
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     361 GAGTTGCGCGTCTCACAGAGCAGCCTCCGGCGGTTGTGGCGCGGTGACGAGAGCGGC   420

Qy     421 GAAGTAGTGCTCCTAGACGCCCCGGTGCTGCTGGGGGGCCCCGAGAGCGGC   480
           ||||||||||||||||||||||||||||||||||||||||||||||||||
Db     421 GAAGTAGTGCTCCTAGACGCCCCGGTGCTGCTGGGGGGCCCCGAGAGCGGC   480

Qy     481 CACGTAGTGTTGCGCTGGCCTGAGACACCCATGACGTCTCACATCCGCTAC   540
           ||||||||||||||||||||||||||||||||||||||||||||||||||
Db     481 CACGTAGTGTTGCGCTGGCCTGAGACACCCATGACGTCTCACATCCGCTAC   540

Qy     541 GAGGTGGACGTCTCGGCGCCGGCAACGCGGAGATCCTGGAG   600
           |||||||||||||||||||||||||||||||||||||||||
Db     541 GAGGTGGACGTCTCGGCGCCGGCAACGCGGAGATCCTGGAG   600

Qy     601 GGCCGCACCGAGTGTGCTGAGCAACCTGCGGGGCGGCCGACACCTTCGGCCGTC   660
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     601 GGCCGCACCGAGTGTGCTGAGCAACCTGCGGGGCGGCCGACACCTTCGGCCGTC   660

Qy     661 CGCGCCGGCTATGGCCTGAGCCGAGCTTCGGCGGCCTTCGGCGGCCGCCTGTG   720
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     661 CGCGCCGGCTATGGCCTGAGCCGAGCTTCGGCGGCCTTCGGCGGCCGCCTGTG   720

Qy     721 TCGCTGCTGACGCCTAGCGACCTGGACCCCTCATCCTGCGTG   780
           ||||||||||||||||||||||||||||||||||||||||||
Db     721 TCGCTGCTGACGCCTAGCGACCTGGACCCCTCATCCTGCGTG   780
```

(left column, rotated)

```
Qy     361 GAGTTGCGCGTCTCACAGAGCAGCCTCCGGCGGTTGTGGCGCGGTGACGAGAGCAATCAAT   420
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     361 GAGTTGCGCGTCTCACAGAGCAGCCTCCGGCGGTTGTGGCGCGGTGACGAGAGCAATCAAT   420

Qy     421 GAAGTAGTGCTCCTAGACGCCCCGGTGGGCGTTGGCGCCGG   480
           |||||||||||||||||||||||||||||||||||||||||
Db     421 GAAGTAGTGCTCCTAGACGCCCCGGTGGGCGTTGGCGCCGG   480

Qy     481 CACGTAGTGTTGCGCTGGCCTGGTGGCAGTGTGACATAGTGGCCATGATGAAGGC   540
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     481 CACGTAGTGTTGCGCTGGCCTGGTGGCAGTGTGACATAGTGGCCATGATGAAGGC   540

Qy     541 TCAGAAGCATCCTCCTGCTCATCTGCTTGCTCTCGGAGAGGGAGGCCTCT   600
           ||||||||||||||||||||||||||||||||||||||||||||||||||
Db     541 TCAGAAGCATCCTCCTGCTCATCTGCTTGCTCTCGGAGAGGGAGGCCTCT   600

Qy     601 GCTGCCAGCTTTGAGTACACATATCCTGACCCCAGCTTCCCAGCTCTTGCCGTCCATGG   660
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     601 GCTGCCAGCTTTGAGTACACATATCCTGACCCCAGCTTCCCAGCTCTTGCCGTCCATAG   660
```

```
RESULT 7
AX008198
LOCUS       AX008198                  1317 bp    DNA          linear       PAT 06-SEP-2000
DEFINITION  Sequence 4 from Patent WO9967360.
ACCESSION   AX008198
VERSION     AX008198.1   GI:9995823
KEYWORDS    .
```

```
Qy  181  GTGTGTTTCTGGGAGGAGCGGCGGAGCGGCTGGGGTGGGCCCGGGCAACTACACAGCTTCTCC  240
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  181  GTGTGTTTCTGGGAGGAGCGGCGGAGCGGCTGGGGTGGGCCCGGGCAACTACACAGCTTCTCC  240

Qy  241  TACCAGCTCGAGGATGAGCCATGGAGCTGTGTCGCCTGCACCAGGCCTCCACGCGTCGT  300
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  241  TACCAGCTCGAGGATGAGCCATGGAGCTGTGTCGCCTGCACCAGGCCTCCACGCGTCGT  300

Qy  301  GGTGCGGTGCGCTTCGTGTTCGTGCTGCTACAGCTGCGACACGTCGAGCTTCGTGCCCCTA  360
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  301  GGTGCCGGTGCGCTTCGTGTTCGTGCTGCTACAGCTGCGACACGTCGAGCTTCGTGCCCCTA  360

Qy  361  GAGTTGCGCGTCACAGACGCCTCCGGCCGCTCCGGCGATATCACCGTGTCATCCACATCAAT  420
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  361  GAGTTGCGCGTCACAGACGCCTCCGGCCGCTCCGGCGATATCACCGTGTCATCCACATCAAT  420

Qy  421  GAAGTAGTGCTCCTAGACGCGCCCGTGGGGCGGTGGTGGCTGACGAGAGCGGC  480
         |||||||||||||||||| | ||||||||||||||||||||||||||||||||
Db  421  GAAGTAGTGCTCCTAGACGCGCCCGTGGGGCATGCGTTGGCTGACGAGAGCGGC  480

Qy  481  CACGTAGTGTTGCGCTGCTGCTCCTGAGACACTGGCATCCGCTCACATCCGCTAC  540
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  481  CACGTAGTGTTGCGCTGCTGCTCCTGAGACACCCATCCGCTCACATCGCTCGCTAC  540

Qy  541  GAGGTGACGTCTCGGCCGGCAACGGCGCAGGAGCGTGGAGGGTGGAGATCCTGGAG  600
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  541  GAGGTGACGTCTCGGCCGGCAACGGCGCAGGAGCGTGGAGGGTGGAGATCCTGGAG  600

Qy  601  GGCCGCACCAGTGTGTGCTGAGCGAGCTTCGGCGCCGCTACACCTTCGCCGTC  660
         |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  601  GGCCGCACCAGTGTGTGCTGAGCGAGCTTCGGCGCCGCTACACCTTCGCCGTC  660

Qy  661  CGCGGCCGGCTATGGCTGAGCCGAGCTTCGGCGGCCCTCGTTGCTGCTGGGTCCTGTG  720
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  661  CGCGGCCGGCTATGGCTGAGCCGAGCTTCGGCGGCCCTCGTTGCTGCTGGGTCCTGTG  720

Qy  721  TCGCTGCTGCTGACGCCTAGCGACCTGGACGCCCCCTCATCCTGACGCCTCATCCTGACGCCTCATCCTGCTG  780
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  721  TCGCTGCTGCTGACGCCTAGCGACCTGGACGCCCCCTCATCCTGACGCCTCATCCTGACGCCTCATCCTGCTG  780
```

```
Qy  781  GTCATCCTGGTGCTGCTGACCGTGCTGCTCCCGCTGGCTCTCCCAGCCGCCGGGGCTCTGAAGCAG  840
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  781  GTCATCCTGGTGCTGCTGACCGTGCTGCTCCCGCTGGCTCTCCCAGCCGCCGGGGCTCTGAAGCAG  840

Qy  841  AAGATCTGGCCTGGCATCCTGGCATCCCGAGCGCCAGAGAGCGAGTTGAAGGCCCTCTTCACCACCCAC  900
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  841  AAGATCTGGCCTGGCATCCTGGCATCCCGAGCGCCAGAGAGCGAGTTGAAGGCCCTCTTCACCACCCAC  900

Qy  901  AAGGGTAACTTCACGGAGGAGACCACCTGCTTCCCTGGAAGTCCTGTGGTGCTGCTGGAGCCCGTGG  960
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  901  AAGGGTAACTTCACGGAGGAGACCACCTGCTTCCCTGGAAGTCCTGTGGTGCTGCTGGAGCCCGTGG  960

Qy  961  ACCCCCTTCACGGAGGAGACCACCTGCGGGGACAGATGATGAGGGCCCCCCGCTGCTGGGG  1020
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  961  ACCCCCTTCACGGAGGAGACCACCTGCGGGGACAGATGATGAGGGCCCCCCGCTGCTGGGG  1020

Qy 1021  ACGATGAGGCAGTGAGCCAGTGGGGACAGATAGGGCCCCCCGCTGCTGGAGCCCAGTGGGC  1080
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1021  ACGATGAGGCAGTGAGCCAGTGGGGACAGATAGGGCCCCCCGCTGCTGGAGCCCAGTGGGC  1080

Qy 1081  AGTGAGCATGCCCAGGATACCTATCTCGGTGTGGACAAATGTTGCTGCTCCGGAACCCG  1140
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1081  AGTGAGCATGCCCAGGATACCTATCTCGGTGTGGACAAATGTTGCTGCTCCGGAACCCG  1140

Qy 1141  CCCAGTGAGGACCTTCCCAGGGCCTGGTCAGTGTGGACATAGTGGCATGGGACATGAAGGC  1200
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1141  CCCAGTGAGGACCTTCCCAGGGCCTGGTCAGTGTGGACATAGTGGCATGGGACATGAAGGC  1200

Qy 1201  TCAGAAGCATCCTCCTGCTCATCGCTTTGGCCTTGGCCTCGAAGCCCCAGAGAGGGAGCCTCT  1260
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1201  TCAGAAGCATCCTCCTGCTCATCGCTTTGGCCTTTGGCCTCGAAGCCCCAGAGAGGGAGCCTCT  1260
```

```
Query Match     85.6%;  Score 1307.4;  DB 6;  Length 1317;
Best Local Similarity   99.5%;  Pred. No. 4.4e-219;
Matches 1311;  Conservative   0;  Mismatches   6;  Indels   0;  Gaps   0;
```

```
Qy    1  ATGGACCACCCCTCGGGGCCGTCCCCTTCTGCGCCCCCAGGTCGCGTCCCTTTGTCTCCTGCT  60
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1  ATGGACCACCCCTCGGGGCCGTCCCCTTCTGCGCCCCCAGGTCGCGTCCCTTTGTCTCCTGCT  60

Qy   61  GGGGCCGCCCTGGGCCGCGCCCCCCGCCTAACCTCCCGGACCCCCAGTTCGAGAGCAAAGCGGCC  120
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61  GGGGCCGCCCTGGGCCGCGCCCCCCGCCTAACCTCCCGGACCCCCAGTTCGAGAGCAAAGCGGCC  120

Qy  121  TTGCTGGCGGCCCCGGGGCCCGGGCCCGGGCCCGAAGAGCTTCGTGCTTCACCGGAGCGGTTGGAGGACTTG  180
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  121  TTGCTGGCGGCCCCGGGGCCCGGGCCCGGGCCCGAAGAGCTTCGTCATCGCTTTGGCCTTTGGCCGGAGCGGTTGGAGGACTTG  180
```

Qy   1261 GCTGCCAGCTTTGAGTACACTATCCTGACCCCAGCTCTTGCGTCCATGG 1317
           ||||||||||||||||||||| ||||||||||||||||||||||||| -
Db   1261 GCTGCCAGCTTTGAGTACACTATCCTGACCACCCCAGCTCTTGCGTCCATAG 1317

RESULT 9
AX008201
LOCUS   AX008201   1317 bp   DNA   linear   PAT 06-SEP-2000
DEFINITION   Sequence 7 from Patent WO9967360.
ACCESSION   AX008201
VERSION   AX008201.1 GI:9995825
KEYWORDS   .
SOURCE   Homo sapiens (human)
  ORGANISM   Homo sapiens
                Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
                Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
  AUTHORS   Matthews,K.E., Bell,D. and Mueller,S.G.
  TITLE   The efficient culture of stem cells for the production of
                hemoglobin
  JOURNAL   Patent: WO 9967360-A 7 29-DEC-1999;
                HEMOSOL INC (CA)
FEATURES   Location/Qualifiers
     source   1..1317
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
     terminator   1315..1317
ORIGIN

Query Match   85.6%; Score 1307.4; DB 6; Length 1317;
Best Local Similarity   99.5%; Pred. No. 4.4e-219;
Matches 1311; Conservative   0; Mismatches   6; Indels   6; Gaps   0;

Qy   1 ATGGACCACCTCGGGGCGTCCTCTGGCCCCAGGTCGGCTCCCTTTGTCTCCTGCTCGCT 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1 ATGGACCACCTCGGGGCGTCCTCTGGCCCCAGGTCGGCTCCCTTTGTCTCCTGCTCGCT 60

Qy   61 GGGGCCGCCTGGCGCCGCCCCCGGGGGCGCCCTAACCTCCCGGACCCCAAGTTCGAGAGCAAAGCGGCC 120
           ||||||||||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||||||
Db   61 GGGGCCGCCTGGCGCCGCCCCCGGGGGCGCCCTAACCTCCCGGACCCCAAGTTCGAGAGCAAAGCGGCC 120

Qy   121 TTGCTGGCGCCCGGGGGCCCGAAGAGCTTCTGTGCTTCACCGAGCGGTTGGAGGACTTG 180
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   121 TTGCTGGCGCCCGGGGGCCCGAAGAGCTTCTGTGCTTCACCGAGCGGTTGGAGGACTTG 180

Qy   181 GTGTGTTTCTGGGAGGAAGCGGCGAGCGGCGGGGTGGGGCCCGGCAACTACAGCTTCTCC 240
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   181 GTGTGTTTCTGGGAGGAAGCGGCGAGCGGCGGGGTGGGGCCCGGCAACTACAGCTTCTCC 240

Qy   241 TACCAGCTCGAGGATGAGCCATGGAAGCTGTGTGCGCCTGACCAGCTCCACGGCTCGT 300
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   241 TACCAGCTCGAGGATGAGCCATGGAAGCTGTGTGCGCCTGACCAGCTCCACGGCTCGT 300

Qy   301 GGTGCGGTGGCCTTCGTGTTCGCCTGCCTACAGCCGACACGTCGAGCTTCGTGCCCCTA 360
           |||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   301 GGTGCCGGTACGCTTCGTGTGTTCGCCTGCCTACAGCCGACACGTCGAGCTTCGTGCCCCTA 360

Qy   361 GAGTTGCGCGTCACAGCAGCCTCCGGCGCCTTCCGGCGCTTCCGGCGCTCCGGCGATATCACCGTCATCCACAAT 420
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   361 GAGTTGCGCGTCACAGCAGCCTCCGGCGCCTTCCGGCGCTTCCGGCGCTCCGGCGATATCACCGTGTCATCCACAAT 420

Qy   421 GAAGTAGTGCTCCTAGACGCGCCCGTGGGGCGGTTGGCGCGGTTGGCATGCTTGGCTGGCATGGCGGC 480
           ||||||||||| |||||||||||||||||||||| |||||||||| |||||||||
Db   421 GAAGTAGTGCTCCCTAGAACGCGCCCGTGGGGCGGTTGGCATGCTTGGCTGGCATGGCGGC 480

Qy   481 CACGTAGTGTTGCGCGTCTGCGCTGGCTCCCGCGGCCCATGAGCGTCTCACATCCGCTAC 540
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   481 CACGTAGTGTTGCGCGTCTGCGCTGGCTCCCGCGGCCCATGAGCGTCTCACATCCGCTAC 540

Qy   541 GAGGTGGACGTCTCGGCGGCAACGGCGCCGGACGTACAGAGGGTAGAGGTGGATCCTGGAG 600
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   541 GAGGTGGACGTCTCGGCGGCAACGGCGCCGGACGTACAGAGGGTAGAGGTGGATCCTGGAG 600

Qy   601 GGCCGCACCGAGTGTGTGCTGAGCAACCTGGCGGGCGGCCGGAGGACGCGGCTACACCTTCGCCGTC 660
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   601 GGCCGCACCGAGTGTGTGCTGAGCAACCTGGCGGGCGGCCGGAGGACGCGGCTACACCTTCGCCGTC 660

Qy   661 CGGCGCGGGTATGGCGTCGTAGGCTTCGGCGTTCGGGCGGCTTCGGGAGCGGCCTGCTGTCGGAGCGCTCGTG 720
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   661 CGGCGCGGGTATGGCGTCGTAGGCTTCGGCGTTCGGGCGGCTTCGGGAGCGGCCTGCTGTCGGAGCGCTCGTG 720

Qy   721 TCGCTGCTGACGCGCTAGCGACCTGACGACCCCCATCCCCTCGAGCTGCTGAGCGTCATCCTCGTG 780
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   721 TCGCTGCTGACGCGCTAGCGACCTGACGACCCCCATCCCCTCGAGCTGCTGAGCGTCATCCTCGTG 780

Qy   781 GTCATCCTGGTGCTGCTGACCGTGCTCGCGCTGCTCCCCACCGGGCTCGCCGGGCTCTGAAGCAG 840
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   781 GTCATCCTGGTGCTGCTGACCGTGCTCGCGCTGCTCCCCACCGGGCGGGGCTCTGAAGCAG 840

Qy   841 AAGATCTGGCCTGGCATCCCGAGCCCCAGAGAGCGAGTTGAAGGCCTCTTCACCACCACCAC 900
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   841 AAGATCTGGCCTGGCATCCCGAGCCCCAGAGAGCGAGTTGAAGGCCTCTTCACCACCACCAC 900

Qy   901 AAGGGTAACTTCCAGTCTGGCTGTACCAGAATGATGGCTGCCTGTGGTGGAGCCCCTGC 960
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   901 AAGGGTAACTTCCAGTCTGGCTGTACCAGAATGATGGCTGCCTGTGGTGGAGCCCCTGC 960

Qy   961 ACCCCCTTCACGGAGGACCACCGTTCCCTGGAAGTCCTTCTCAGAGCGCTGTGGGG 1020
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   961 ACCCCCTTCACGGAGGACCACCGTTCCCTGGAAGTCCTTCTCAGAGCGCTGTGGGG 1020

Qy   1021 ACGATGCAGGCAGTGAGCCGGGACAGATGATGAGGCCCCTGCTTGGAGCCAGTGGC 1080
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1021 ACGATGCAGGCAGTGAGCCGGGACAGATGATGAGGCCCCTGCTTGGAGCCAGTGGC 1080

Qy   1081 AGTGAGCATGCCCAGATACCTATCTGGTCGTGCTGGACAAATGGTTGTGCGCCCCGGACCCG 1140
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1081 AGTGAGCATGCCCAGATACCTATCTGGTCGTGCTGGACAAATGGTTGTGCGCCCCGGACACCG 1140

Qy   1141 CCCAGTGAGGACCTCCCAGGGCCTGGTCAGTGTGGACATAGTGGCATGGTGAAGGC 1200
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1141 CCCAGTGAGGACCTCCCAGGGCCTGGTCAGTGTGGACATAGTGGCATGGTGAAGGC 1200

Qy   1201 TCAGAAGCATCTCCTGTGCTTCATCGTCTTTGGCCTCGAAGCCCAGAGGAGGAGCCTCT 1260
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1201 TCAGAAGCATCTCCTGTGCTTCATCGTCTTTGGCCTCGAAGCCCAGAGGAGGAGCCTCT 1260

Qy   1261 GCTGCCAGCTTGAGTACACTATCCTGGACCCCCAGCTCTTGCGTCCATGG 1317
           ||||||||||||||||||||||||||||||||||||||||||||||| -
Db   1261 GCTGCCAGCTTGAGTACACTATCCTGGACCCCCAGCTCTTGCGTCCATAG 1317

RESULT 10
BC019092
LOCUS   BC019092   2154 bp   mRNA   linear   PRI 02-JAN-2004
DEFINITION   Homo sapiens erythropoietin receptor, mRNA (cDNA clone
                IMAGE:4894283), partial cds.
ACCESSION   BC019092
VERSION   BC019092.2 GI:40226074
KEYWORDS   .
SOURCE   Homo sapiens (human)
  ORGANISM   Homo sapiens
                Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
                Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 2154)
  AUTHORS   Strausberg,R.L., Feingold,E.A., Grouse,L.H., Derge,J.G.,
                Klausner,R.D., Collins,F.S., Wagner,L., Shenmen,C.M., Schuler,G.D.,
                Altschul,S.F., Zeeberg,B., Buetow,K.H., Schaefer,C.F., Bhat,N.K.,
                Hopkins,R.F., Jordan,H., Moore,T., Max,S.I., Wang,J., Hsieh,F.,
                Diatchenko,L., Marusina,K., Farmer,A.A., Rubin,G.M., Hong,L.,
                Stapleton,M., Soares,M.B., Bonaldo,M.F., Casavant,T.L.,
                Scheetz,T.E., Brownstein,M.J., Usdin,T.B., Toshiyuki,S.,
                Carninci,P., Prange,C., Raha,S.S., Loqueliano,N.A., Peters,G.J.,
                Abramson,R.D., Mullahy,S.J., Bosak,S.A., McEwan,P.J.,
                McKernan,K.J., Malek,J.A., Gunaratne,P.H., Richards,S.,

Worley,K.C., Hale,S., Garcia,A.M., Gay,L.J., Hulyk,S.W.,
Villalon,D.K., Muzny,D.M., Sodergren,E.J., Lu,X., Gibbs,R.A.,
Fahey,J., Helton,E., Ketteman,M., Madan,A., Rodrigues,S.,
Sanchez,A., Whiting,M., Madan,A., Young,A.C., Shevchenko,Y.,
Bouffard,G.G., Blakesley,R.W., Touchman,J.W., Green,E.D.,
Dickson,M.C., Rodriguez,A.C., Grimwood,J., Schmutz,J., Myers,R.M.,
Butterfield,Y.S., Krzywinski,M.I., Skalska,U., Smailus,D.E.,
Schnerch,A., Schein,J.E., Jones,S.J. and Marra,M.A.

TITLE   Generation and initial analysis of more than 15,000 full-length
             human and mouse cDNA sequences
JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
PUBMED   12477932
REFERENCE   2 (bases 1 to 2154)
AUTHORS   Strausberg,R.
TITLE   Direct Submission
JOURNAL   Submitted (07-DEC-2001) National Institutes of Health, Mammalian
             Gene Collection (MGC) Cancer Genomics Office, National Cancer
             Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
             USA
REMARK   NIH-MGC Project URL: http://mgc.nci.nih.gov
COMMENT   On Dec 19, 2003 this sequence version replaced gi:19387983.
             Contact: MGC help desk
             Email: cgapbs-r@mail.nih.gov
             Tissue Procurement: ATCC
             cDNA Library Preparation: Rubin Laboratory
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: National Institutes of Health Intramural
             Sequencing Center (NISC),
             Gaithersburg, Maryland;
             Web site: http://www.nisc.nih.gov/
             Contact: nisc_mgc@nhgri.nih.gov
             Akhter,N., Ayele,K., Beckstrom-Sternberg,S.M., Benjamin,B.,
             Blakesley,R.W., Bouffard,G.G., Breen,K., Brinkley,C., Brooks,S.,
             Dietrich,N.L., Granite,S., Guan,X., Gupta,J., Haghighi,P.,
             Hansen,N., Ho,S.-L., Karlins,E., Kwong,P., Laric,P., Legaspi,R.,
             Maduro,Q.L., Masiello,C., Maskeri,B., Mastrian,S.D., McCloskey,J.C.,
             McDowell,J., Pearson,R., Stantripop,S., Thomas,P.J., Touchman,J.W.,
             Tsurgeon,C., Vogt,J.L., Walker,M.A., Wetherby,K.D., Wiggins,L.,
             Young,A., Zhang,L.-H. and Green,E.D.

             Clone distribution: MGC clone distribution information can be found
             through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
             Series: IRAL Plate: 39 Row: j Column: 20
             This clone was selected for full length sequencing because it
             passed the following selection criteria: matched mRNA gi: 4557561.
FEATURES   Location/Qualifiers
     source   1..2154
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:4884283"
                     /tissue_type="Brain, neuroblastoma"
                     /clone_lib="NIH_MGC_19"
                     /lab_host="DH10B-R"
                     /note="Vector: pOTB7"
ORIGIN

Query Match   83.6%; Score 1276.6; DB 9; Length 2154;
Best Local Similarity   99.7%; Pred. No. 9.9e-214;
Matches 1279; Conservative   0; Mismatches   4; Indels   0; Gaps   0;

Qy   245 AGCTCGAGGATGAGCCATGAGAAGCTGTGTCGCCTGACCAGGCTCCGTCGTG 304
                |||||||||||||||||||||||||||||||||||||||||||||||||||
Db   630 ATCGAGGGGATGAGCCATGAGAAGCTGTGTCGCCTGACCAGGCTCCGTCGTGTG 689

Qy   305 CGGTGCGCTTCGTGTGTTCGCTGCTCAGCACACGTCGAGCTTCCGTGCCCCTAGAGT 364
                |||||||||||||||||||||||||||||||||||||||||||||||||||
Db   690 CGGTGCGCTTCGTGTGTTCGCTGCTCAGCACACGTCGAGCTTCCGTGCCCCTAGAGT 749

Qy   365 TGCGCGGTCACAGCAGCCTCCGGCGCGCTCCGGACCATCACCGTGTCATCCACATCAATGAAG 424
                ||||||||||||||||||||||||||||||||||||||||||||||||||
Db   750 TGCGCGGTCACAGCAGCCTCCGGCGCGCTCCGGACCATCACCGTGTCATCCACATCAATGAAG 809

Qy   425 TAGTGCTCCTAGACGCCCCCGTGGGGCTGGTCGTGGCGGTTGGCTGAGAGAGCGGCCACG 484
                |||||||||||||||||||||||||||||||||||||||||||||||||||
Db   810 TAGTGCTCCTAGACGCCCCCGTGGGGCTGGTCGTGGCGGTTGGCTGAGAGAGCGGCCACG 869

Qy   485 TAGTGTTGCGCTGCTCCCGCGCCGCCTGAGACACCCATGAGACGTCTCACATCCGCTACGAGG 544
                |||||||||||||||||||||||||||||||||||||||||||||||||||
Db   870 TAGTGTTGCGCTGCTCCCGCGCCGCCTGAGACACCCATGAGACGTCTCACATCCGCTACGAGG 929

Qy   545 TGGACGTCTCGGCCGGCAACGGGCAGGAGGGTACAGAGGGTGGAGATCCTGGAGGGCC 604
                |||||||||||||||||||||||||||||||||||||||||||||||||||
Db   930 TGGACGTCTCGGCCGGCAACGGGCAGGAGCGTACAGAGGGTGGAGATCCTGGAGGGCC 989

Qy   605 GCACCGAGTGTGTGCTGAGCAACCTGGGGGCCCGGGGGCGCTACACCTTCGCGTCCGG 664
                |||||||||||||||||||||||||||||||||||||||||||||||||||
Db   990 GCACCGAGTGTGTGCTGAGCAACCTGGGGGCCCGGGGGCGCTACACCTTCGCGTCCGCG 1049

Qy   665 CGGGTATGCTGAGCCGAGCTTCGGCGGCTTCGGCGGCCCTCTGGAGGCGTGTGTGCGC 724
                |||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1050 CGCGTATGCTGAGCCGAGCTTCGGCGGCTTCGGCGGCCCTCTGGAGGCGTGTGTGCGC 1109

Qy   725 TGCTGACGGCTAGGACCTGGACCCCCATCCTGACGCTCTCCTCATCCTGGTCA 784
                |||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1110 TGCTGACGGCCTAGGACCTGGACCCCCATCCTGACGCTCTCCTCATCCTGGTCA 1169

Qy   785 TCCTGGTGCTGCTGACCGTGCTCTCCCACCGCCGGGGCCTCTTCACCACCCACAAGG 844
                |||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1170 TCCTGGTGCTGCTGACCGTGCTCTCCCACCGCCGGGGCCTCTTCACCACCCACAAGA 1229

Qy   845 TCTGGCCTGGCATCCCAGAGCCCGGGGACAGAGGAGCGAGTTGAAGGCTTGAAGGCCTTGAAGGA 904
                |||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1230 TCTGGCCTGGCATCCCAGAGCCCGGGGACAGAGGAGCGAGTTTGAAGGCTTTGAAGGAGA 1289

Qy   905 GTAACTTCCAGCTGTGGCTGTACCAGGCGTGCTCGTGTGGGGAGCCCTGCACCC 964
                |||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1290 GTAACTTCCAGCTGTGGCTGTACCAGGCGTGCTCGTGTGGGGAGCCCTGCACCC 1349

Qy   965 CCTTCACGGAGGACACCCACCTGCTTCCTGGAGTCCTTCCTGCTGCTGGGGGACGA 1024
                |||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1350 CCTTCACGGAGGACACCCACCTGCTTCCTGGAGTCCTTCCTGCTGCTGGGGGACGA 1409

Qy   1025 TGCAGGCAGTGAGCGGGACAGATGAGGAGGCCCCGCTGGAGCCAGTGGCAGTG 1084
                |||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1410 TGCAGGCAGTGAGCGGGACAGATGAGGAGGCCCCGCTGGAGCCAGTGGCAGTG 1469

Qy   1085 AGCATGCCCAGGATACCTATCGTGTGCTGGACAAATGGTTGCTGCCCCGGAACCCGCCCA 1144
                |||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1470 AGCATGCCCAGGATACCTATCGTGTGCTGGACAAATGGTTGCTGCCCCGGAACCCGCCCA 1529

Qy   1145 GTGAGGACCTCCCAGGGACCTGGCAGTGTGACATAGTGACCAGGCTCAG 1204
                |||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1530 GTGAGGACCTCCCAGGGACCTGGCAGTGTGACATAGTGACCAGGCTCAG 1589

Qy   1205 AAGCATCTCCTGCTACATCCTGCTCGTCTTTGGCTCTGCCGAAGCCCAGAGAGGAGCCTCTGCTG 1254
                |||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1590 AAGCATCTCCTGCTACATCCTGCTCGTCTTTGGCTCTGCCGAAGCCCAGAGAGGAGCCTCTGCTG 1649

Qy   1265 CCAGCTTTGAGTACACATATCCTGGACCCCAGCTCTTCGCTCCATGGACACTGT 1324
                |||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1650 CCAGCTTTGAGTACACATATCCTGGACCCCAGCTCTTCGCTCCATGGACACTGT 1709

Qy   1325 GCCCTGAGCTGCCCCCTACCCCACCTAAAGTACCTGTACCTGTGGTATCTGACT 1384
                |||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1710 GCCCTGAGCTGCCCCCTACCCCACCTAAAGTACCTGTACCTGTGGTATCTGACT 1769

Qy   1385 CTGGCATCCAACTGACTACAGGTCACACCTCAGGGAGCTTCCGTGCCCCTAGAGT 1444
                |||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1770 CTGGCATCCAACTGACTACAGCTCACACCTCAGGGAGCTTCCGTGCCCCTAGAGT 1829

Qy   1445 ATGGCCCCTACTTCAACCCTTATGAGAACAGCCTTATCCCACGTGTCATCCACATCAATGAAG 1504
                |||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1830 ATGGCCCCTACTTCAACCCTTATGAGAACAGCCTTATCCCACGTGTCATCCACATCAATGAAG 1889

Qy   1505 CCAGCTATGTGGCTTGCTTCTTAG 1527

RESULT 11
AK074082

LOCUS	AK074082	5761 bp	mRNA	linear	PRI 15-FEB-2002
DEFINITION	Homo sapiens mRNA for FLJ00153 protein.				
ACCESSION	AK074082				
VERSION	AK074082.1 GI:18676511				
KEYWORDS	fis (full insert sequence).				
SOURCE	Homo sapiens (human)				
ORGANISM	Homo sapiens				

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE   1
  AUTHORS   Jikuya,H., Takano,J., Nomura,N., Kikuno,R., Nagase,T. and Ohara,O.
  TITLE   The nucleotide sequence of a long cDNA clone isolated from human
            spleen
  JOURNAL   Published Only in Database (2002)
REFERENCE   2 (bases 1 to 5761)
  AUTHORS   Jikuya,H., Takano,J., Nomura,N., Kikuno,R., Nagase,T. and Ohara,O.
  TITLE   Direct Submission
  JOURNAL   Submitted (21-JAN-2002) Takahiro Nagase, Kazusa DNA Research
            Institute, Department of Human Gene Research; 1532-3, Yana,
            Kisarazu, Chiba 292-0812, Japan (E-mail:cdnaintc@kazusa.or.jp,
            URL:http://www.kazusa.or.jp/NEDO, Tel:81-438-52-3913,
            Fax:81-438-52-3914)
COMMENT   NEDO human cDNA sequencing project supported by Ministry of
            Economy, Trade and Industry of Japan; cDNA full insert and 5'- &
            3'-end one pass sequencing: Research Association for Biotechnology;
            cDNA library construction and clone selection: Kazusa DNA Research
            Institute.
FEATURES   Location/Qualifiers
     source   1..5761
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="FLJ00153"
                     /tissue_type="spleen"
                     /dev_stage="adult"
                     /note="vector:pBluescriptII SK plus"
     gene   1..5761
                     /gene="FLJ00153"
     CDS   <595..2088
                     /gene="FLJ00153"
                     /note="Start codon is not identified."
                     /codon_start=1
                     /evidence=not_experimental
                     /product="FLJ00153 protein"
                     /protein_id="BAB84908.1"
                     /db_xref="GI:18676512"
                     /translation="ARAGRGAASGPPRVAQTSDPDSSEACAEEEEGLMPQGPQLLDFS
                     VDEVAEQLTLIDLELFSKVRLYECLGSVWSQRDRPGAAGASPTVRATVAQFNTVTGCV
                     LGSVLGAPGLAAPQRAQRLEKWIRIAQRCRELRNFSSLRAILSALQSNPIYRLKRSWG
                     AVSREPLSTFRKLSQIFSDENNHLSSREILFQEEATEGSQEEDNTPGSLPSKPPPGPV
                     PYLGTFLTDLVMLDTALPDMLEGDLINFEKRRKEWEILARIQQLQRRCQSYTLSPHPP
                     ILAALHAQNQLTEEQSYRLSRVIEPPAASCPSSPIRRRISLTKRLSAKLAREBKSSSP
                     SGSPGDPSSPTSSVSPGSPPSSPPSRDAPAGSPPASPGPGPSTKLPLSLDLPSPRPF
                     ALPLGSPRIPLPAAQSSEACVIRVSIDNDHGNLYRSILLTSQDKAPSVVRRALQKHNV
                     PQPWACDYQLPQVLPFGDRGWEEMJISAIKSVGTFGFMFAMCQARFYR"

ORIGIN

Query Match   83.6%; Score 1276.6; DB 9; Length 5761;
Best Local Similarity   99.7%; Pred. No. 8.2e-214;
Matches 1279;   Conservative   0; Mismatches   4;   Indels   0;   Gaps   0;

Qy   245 AGCTCGAGGATGAGCCATGAGCAGCTGTGTCGCCTGACACCAGGCTCCCACGGCTCGGTG 304
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3702 ATCGAGGGGATGAGCCATGAGCAGCTGTGTCGCCTGACACCAGGCTGTCCCACGGCTCGGTG 3761

Qy   305 CGGTGCGCTTCTGGTGTTCGCTGCCTACAGCCGACACGTCGGCTTCGTGCCCCCTAGAGT 364
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4842 CTGGCATCTCAACTGACTACAGCTCAGCGGGACTCCCACGGCTCCCACGGGTCGTGGTG 4901

Db	1890	CCAGCTATGTGGCTTGCTCTTAG			1912
Db	3762	CGGTGCGCTTCTGGTGTTCGCTGCCTACAGCCGACACGTCGGCTTCGTGCCCCCTAGAGT			3821
Qy	365	TGCGCGGTCACAGCAGCCTCCGGCGATATCACCGTGCATCCACATCAATGAAG			424
Db	3822	TGCGCGGTCACAGCAGCCTCCGGCGATATCACCGTGCATCCACATCAATGAAG			3881
Qy	425	TAGTGCTTCCTAGACGCCCCGTGGGCTGGTCGGCCGGTTGGCTGGTGGCCACG			484
Db	3882	TAGTGCTTCCTAGACGCCCCGTGGGCGGTGGGCCGGTTGGCTGGTGGGCGGCCACG			3941
Qy	485	TAGTGTTGCGCTGGCTCGGCCTGAGACACCATGACGTCTCACATCCGCTACGAGG			544
Db	3942	TAGTGTTGCGCTGGCTCGGCCTGAGACACCATGACGTCTCACATCCGCTACGAGG			4001
Qy	545	TGGACGTCTCGGCCGGCAACGCGCAGGAGCGTACAGAGGGTGGAGATCCTGGAGGGCC			604
Db	4002	TGGACGTCTCGGCCGGCAACGCGCAGGAGCGTACAGAGGGTGGAGATCCTGGAGGGCC			4061
Qy	605	GCACCGAGTGTGCTGAGCAACCTGCGGGGCCGGCAACCTGCGCTACACCTTCGCGCTCCGCG			664
Db	4062	GCACCGAGTGTGCTGAGCAACCTGCGGGGCCGGCAACCTGCGCTACACCTTCGCGCTCCGCG			4121
Qy	665	CGGGTATGGCTGAGCGGAGCTTCGGCGGCTTCGGAGCTTCGGAGCTTCGGAGCTTCGGAGC			724
Db	4122	CGGGTATGGCTGAGCGGAGCTTCGGCGGCTTCGGAGCTTCGGAGCTTCGGAGC			4181
Qy	725	TGCTGACGCGTAGCGACCTGGAACCCCCTCATCCTGACGCTCTCCTCATCCTCGTGGTCA			784
Db	4182	TGCTGACGCGTAGCGACCTGGAACCCCCTCATCCTGACGCTCTCCTCATCCTCGTGGTCA			4241
Qy	785	TCCTGGTGCTGCTGACGCTGTCGCGGTCGCGGCTCTGAAGCAGAGA			844
Db	4242	TCCTGGTGCTGCTGACGCTGTCGCGGTCGCGGCTCTGAAGCAGAGA			4301
Qy	845	TCTGGCCTGGCATCCGAGCCCAGAGAGCGAGTTGAAGGCCTTCACCACCCACAAGG			904
Db	4302	TCTGGCCTGGCATCCGAGCCCAGAGAGCGAGTTTGAAGGCCTTCACCACCCACAAGG			4361
Qy	905	GTAACTTCCAGTGTGCTGTACCAGATGATGGCTCGCCTGTGGTGGAGCCCCTGCACCC			964
Db	4362	GTAACTTCCAGTGTGCTGTACCAGAATGATGGCTCGCCTGTGGTGGAGCCCCTGCACCC			4421
Qy	965	CCTTCACGGAGGACCACCTGCTTCCTGGAAGTCCTCTCACAGAGCGCTGCTGGGGGACGA			1024
Db	4422	CCTTCACGGAGGACCACCTGCTTCCTGGAAGTCCTCTCACAGAGCGCTGCTGGGGGACGA			4481
Qy	1025	TGCAGGCAGTGGAGCCCGGGACAGATGAGAGGCCCCGTGCTGGAGCGCAGTGGCAGTG			1084
Db	4482	TGCAGGCAGTGGAGCCGGGGACAGATGAGAGGCCCCGTGCTGGAGCGCAGTGGCAGTG			4541
Qy	1085	AGCATGCCAGGATACCTATCTGGTGGACAAATGTTGCTCCCCGGAACCGCGCCCA			1144
Db	4542	AGCATGCCAGGATACCTATCTGGTGGACAAATGTTGCTCCCCGGAACCGCGCCCA			4601
Qy	1145	GTGAGGACCTCCCAGGGCCTGGCAGTGTGGACATAGTGGGACCATGATGAGAGGCTCAG			1204
Db	4602	GTGAGGACCTCCCAGGGCCTGGCAGTGTGGACATAGTGGGACCATGATGAGAGGCTCAG			4661
Qy	1205	AAGCATCTCCTGCTCTGCTTTGGCCTCGAAGCCAGCCAGCCCAGAGGAGGAGCCTCTGCGTG			1264
Db	4662	AAGCATCTCCTGCTCTCGCTTTGGCCTCGAAGCCAGCCAGCCCAGAGGAGGAGCCTCTGCGTG			4721
Qy	1265	CCAGCTTTGAGTACACTATCCTGGACCCCCAGCTCTTCCATGGACACTGT			1324
Db	4722	CCAGCTTTGAGTACACTATCCTGGACCCCCAGCTCTTCCATGGACACTGT			4781
Qy	1325	GCCCTGAGCTGCGCCCCTACCCCACCCCACCCCACCCCACCTAAAGTACCTGTACCTTGTGGTATCTGACT			1384
Db	4782	GCCCTGAGCTGCGCCCCTACCCCACCCCACCCCACCTAAAGTACCTGTACCTTGTGGTATCTGACT			4841
Qy	1385	CTGGCATCTCAACTGACTACAGCTCAGCGGGACTCCCAGGAGCCCAAGGGGCTTATCCG			1444
Db	4842	CTGGCATCTCAACTGACTACAGCTCAGCGGGACTCCCAGGAGCCCAAGGGGCTTATCCG			4901

```
Qy   1445 ATGCCCCTACTCCAACCCTTATGAGAACAGCCTTATCCCAGCCGCTGAGCCTCTGCCCC 1504
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4902 ATGGCCCCTACTCCAACCCTTATGAGAACAGCCTTATCCCAGCCGCTGAGCCTCTGCCCC 4961

Qy   1505 CCAGCTATGTGGCTTGCTCTTAG 1527
          |||||||||||||||||||||||
Db   4962 CCAGCTATGTGGCTTGCTCTTAG 4984

RESULT 12
AF274305
LOCUS       AF274305       1843 bp    mRNA         linear   MAM 08-MAY-2003
DEFINITION  Sus scrofa erythropoietin receptor mRNA, complete cds.
ACCESSION   AF274305
VERSION     AF274305.1  GI:8572251
KEYWORDS    .
SOURCE      Sus scrofa (pig)
  ORGANISM  Sus scrofa
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
REFERENCE   1  (bases 1 to 1843)
  AUTHORS   Pearson,P.L., Smith,T.P.L., Sonstegard,T.S., Klemcke,H.G.,
            Christenson,R.K. and Vallet,J.L.
  TITLE     Porcine erythropoietin receptor: molecular cloning and expression
            in embryonic and fetal liver
  JOURNAL   Domest. Anim. Endocrinol. 19 (1), 25-38 (2000)
  MEDLINE   20419811
   PUBMED   10962196
REFERENCE   2  (bases 1 to 1843)
  AUTHORS   Pearson,P.L., Smith,T.P.L., Sonstegard,T.S., Klemcke,H.G.,
            Christenson,R.K. and Vallet,J.L.
  TITLE     Direct Submission
  JOURNAL   Submitted (01-JUN-2000) Reproduction Unit, Roman L. Hruska, U.S.
            Meat Animal Research Center, P.O. Box 166, Clay Center, NE 68933,
            USA
FEATURES             Location/Qualifiers
     source          1..1843
                     /organism="Sus scrofa"
                     /mol_type="mRNA"
                     /db_xref="taxon:9823"
                     /tissue_type="liver"
                     /dev_stage="fetus"
     CDS             160..1689
                     /note="cytokine receptor"
                     /codon_start=1
                     /product="erythropoietin receptor"
                     /protein_id="AAF77065.1"
                     /db_xref="GI:8572252"
                     /translation="MYHFGATLWPGVGSLCLLLAGATWAPSPNSPDAKFESKAALLAA
                     RGPEELLCFTERLEDLVCFWEEAGSAGVGPEDYSFSYQLEGEPWKPCHLHQGPTARGS
                     VRFWCSLPTADTSSFVPLELRVTEVSSGAPRYHRIIHNEVVLLDPPAGLLARRAEES
                     GHVVLRWLPPPGAPMASLIRYEVNISTENAAGGVQRVEILDGRTECVLSNLRGGTRYT
                     FMVRARMAEPSFGGFWSAWSEPASLLTASDLDPLILTLSLILVLILLLAVLALLSHR
                     RTLKQKIWPGIPSPEGEFEGLFTTHKGNFQLWLYQTDGCLWWSPCTPFAEDPPAPLEV
                     LSERCWGVTQAVEPADDEGSLLEPVGSEHARDTYLVLDKWLLPRRPASEDLPQPGGD
                     LDMAAMDEASEASFCSSALALKPGPEGASAASFEYTILDPSSQLLRPRALPAELPPTP
                     PHLKYLYLVVSDSGISTDYSSGSGEQTQGGSSSGPYSNPYENSLVPAPEPSPPNYVTC
                     S"
ORIGIN
Query Match      75.7%;  Score 1155.6;  DB 4;  Length 1843;
Best Local Similarity   85.2%;  Pred. No. 1.6e-192;
Matches 1303;  Conservative   0;  Mismatches 224;  Indels  3;  Gaps  1;

Qy      1 ATGGACCACCTCGGGGCGCCGGGGCGCCCTCCTGCCCCAGGTCGCCTCCCTTTGTCTCTGCGCT 60
          |||  ||||||   |||||||||||||||||||||||||||||  ||||||| ||||||||
Db    160 ATGTATCATTTCGGGGCGCCAACCCTCTGGCCAGGAGTTGGCTCCTCCTGTCTCTTCGCTAGCT 219

Qy     61 GGGGCCCGCTGGGCGCGCGCGCCCGGCCTAACCTCCCGGACCCCCAAGTTCGAGAGCAAAGCGGCC 120
          |||||  |||||||||||| ||| ||| |  |||||||||||||||||||||||||||||||||
Db    220 GGGGGCCACCTGGGCTCCTTCACCGAACTGCTTCACCCCGACGCGCCCGAGCAAGTTTGAAAGCAAAGCGGCC 279
```

```
Qy    121 TTGCTGGGCGGGCCCGGGGGCCCGGAAGAGCTTCTGTGCTTCACGGAGCGGTTGGAGGACTTG 180
          |||||| ||||||||||||||||||||||||||||||||||||| ||||||||||||||||||
Db    280 CTGTTGGGCCAGCCCGGCGGGCCTGAAGAGCTTCTGTGCTTCACGGAGCGGTTGGAGGACTTG 339

Qy    181 GTGTGTTTCTGGGAGGAGGCGGGCAGCGCTGGGGTGGGCCCGGCCAACTACAGCTTCTCC 240
          ||| || ||||||||||||| || ||||||||||||||||||||||||||||||||||||
Db    340 GTGTGCTTCTGGGAGGAGGCGGGAAGCGCTGGAGTGGGCCCAGCCAGAGACTACAGCTTCTCC 399

Qy    241 TACCAGCTCGAGATGAGCCATGAAGTCGTGTCGCTGCTCACCAGCCTGCTCAAGGGCCCGCGT 300
          ||||||||| |||||||| |||| ||||||||||||||||||| |||||||||||||||||||
Db    400 TACCAGCTTGAGGTGAGCCGTGCCGTGTCCATGCCGCTGCTCGATGCCCACCCGCTGCCCGCCGC 459

Qy    301 GGTGCGGTGCGCTTCTGGTGTTCGCTGCCTCACAGCCGACACGTCGAGCTTCGTGCCCCTA 360
          |||||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||
Db    460 GGCTCGGTGCGTTTCTGGTGTTCGCTGCCTCGCTGCCGCACAGTCGAGCTTCGTGCCCCTA 519

Qy    361 GAGTTGCGCGTCACAGCAG---CCTCCGCGCTTCCGGCGATATCACCGTGTCATCACATC 417
          |||||||||||||||||||   ||||||||||||||||||||| |||||||||||||||||
Db    520 GAGCTGCGCGTCACTGAGGTTTCCTCGGGCGTTCCTGGGCGTTCCACCACCGTATCATCCACATC 579

Qy    418 AATGAAGTGCTCCTAGACGCCCCGTGGGGCGCGGGTTGGCGTGACGAGAGC 477
          ||  ||||||||||||||||||||||||||||||||||||||||||||||||
Db    580 AACGAAGTGGTGCTGCTAGACCCTCCGGCCGGGCTGCTGGCACGGCGGGCCGAAGAGAGT 639

Qy    478 GGCCACGTAGTGTTGCGCTGGCTCCCGCCTGAGACACCCATGAGCGTCTCACATCCGC 537
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    640 GGCCACGTGGTACTGCGGTGGCTGCCGCCGCCGGGGGCACCCATGGCCAGCCTTATCCGC 699

Qy    538 TACGAGGTTGGACGTCTCGGCGCCGGCAACGGCGTACGAGGGTGGAGATCCTG 597
          |||||||| ||||| |||||||||||||||||||  ||||  ||||||||||||
Db    700 TATGAGGTGAACATCTCGACAGAGAACGCCGCAGGAGGCGTGCAGAGGGTGGAAATCCTC 759

Qy    598 GAGGGCGCACCGAGTGTGTGCTGCACCTGCTGAGCAACCTTCGCGGCCGC 657
          |||||||||||||||||||||||||||||||||||||||||||||||||
Db    760 GACGGGCGCACTGAGTGCGTGCTGAGCAACCTGCGGGGCGGGACCCGGTACACCTTCATG 819

Qy    658 GTCCGCGCGCGTATGGCTGACGCCGAGCTTCGGAGGCCTTCTGGAGCGCCTGGAGCCT 717
          ||| |||||||||| ||||| ||||||| ||||||||||||||||| ||||||||||||
Db    820 GTACGCGCGCGAATGGCAGAGCCCAGCTTCGGCGGCTTCTGGAGCGCCTGGTCCGAGCCT 879

Qy    718 GTGTCGTGCTGACGCGTAGGCAGCTTGACCCCTCATCCTCATCCCTCATCCTC 777
          |||||||||| ||||| ||||||||||| |||||||||||||||||||||||
Db    880 GCGTCGTGCTGCAGCTAGTGACTTGGACGTGTCTGCTCTCCCGCTGCCCCACCGCGGGCTCTGAAG 939

Qy    778 GTGGTCATCCTGTGCTTGCTGACGTGCTGCGCCGCCGGGCTCTGAAG 837
          |||| |||| ||||||| ||| |||||||||||||||| ||  |||||
Db    940 GTGCTCATCTCGCTGCTCTGCTGCCGTGCTGCTGCGCCCGTCGGACTCTGAAG 999

Qy    838 CAGAAGATCTGGCCTGCATCCGAGCCCAGAGAGCAGTTGAAGGCCTTCACCACC 897
          ||||||||||||||| ||||||||||| |||||| |||||||||| ||||||||
Db   1000 CAGAAGATCTGGCCCTGTATCCCAAGCCCTGAGGGAGTTTGAGGGTTCTCTTCACACC 1059

Qy    898 CACAAGGGTAACTTGCCCAGGGGCGTGCTACCAGAGATGAATGGCTGAGCCCC 957
          ||||||| |||| |||||| |||||||||||||| || || |||||||||||
Db   1060 CACAAGGGTAATTTCCAGCTGTGGCTGTACCAAACTGATGGTTGTGTGGAGCCCCC 1119

Qy    958 TGCACCCCCTACGGAGGACCTGTTCCTGGAAGTCTTCCTCAGAGCGGCTGCTGG 1017
          ||||||||| ||||||||| |||| |||||||||| ||||||||||||||||||
Db   1120 TGCACCCCCTTTTGCAGAGGACCCACCCGCCCCTGCCGAAGACCCTCCTGCTCCG 1179

Qy   1018 GGGACAGATGCAGGCAGTGGAGCGGGACAGATGATGAGGGCCCCCTGCTGCCAGTG 1077
          ||||| |||||||| |||||||||||||| ||||| |||||||||||||||||
Db   1180 GGGGTGACCAAGCAGTGGAACGACGGGGTTCCCTGTGCCGCGGGGGCCCGGAAC 1239

Qy   1078 GGCAGTGAGCATGCCAGGATACCTATCTGGTGTCGGACAAATGGTTGCCCCGGAAC 1137
          |||||||| ||||| |||||||| || ||||||| ||||| ||||||||||||
Db   1240 GGCAGTGAACATGCCAGGATACCTACCTGGTGCTGGACAAGTGGCTGTGCTCCCCGAGG 1299

Qy   1138 CCGCCCAGTGAGGACCTTCCAGGGCCTAACCTCCCGGACCCCAAGTTCGAGAGCAAAGCGGCC 1197
          ||||||||||||||||| |||||||| ||||| ||||||||||||| ||||||||||||||
Db   1300 CCGGCCAGTGAGGACCTCCCAGAGCCTGGTGGTGATATTGGACATGGCCGCCATGGATGAA 1359

Qy   1198 GGCTCAGAAGCATCCTCCTGCTCATCGTCTTGGCCTCGAAGACGCCCAGCCAGGGGAGCC 1257
          ||||| |||||||||||||||||||||| ||||||| |||||||||||||||||||||
Db   1198
```

```
Qy      1 ATGGACCACCTCGGGGCGTCCCTCTGCGCTCGGGCTCGGCTCCTCCCTTTGTCTCCTGCTCGCT  60
          ||||||||||||||||||||||| |||||||||||||| ||||| |||  |||  ||| |||
Db     28 ATGGACAAACTCAGGGTGCCCCTCCTGCCCCGTGTCCCTCGGGTAGGCCCCCTCGTCTCTACTTGCT  87

Qy     61 GGGGCGCCTGGGCGGCCCCCGGCTAACCTCCCCGGACCCCAAGTTCGAGAGCAAAGCGGCC  120
          ||||| ||||||  |||  ||| ||||||||  |||||||||||| ||||||| ||||| |||
Db     88 GGGGCAGCCGCTGGGCCACCTCACCCCACCCCCAACCCCCAAGTTTGAGAGCAAAGCGGCC  147

Qy    121 TTGCTGGCGGCCCGGGGCCCGAAAGAGTTCTGTGTTCACCGAGCGGTTGGAGGACTTG  180
          ||||| |||||| ||||||||||||||  |||||||||  |||||| |||||||||||||
Db    148 CTGCTCTGGCATCCCGGGCTCCGAAGAACTTCTGTGCTCACCCACCAACCGTTGGAAGACTTG  207

Qy    181 GTGTGTTTCTGGAGGAAGCGGCGAGCGGCGAGCTCGGGGTGGGCCAACTACAGCTTCTCC  240
          |||||||| ||||||||| |||| ||||||| ||||| ||| |||||||||||||||||
Db    208 GTGTGTTTCTGGAGGGTGAGTCACGAAAGTCATGTGAGCCGACACCGTGACCAGGCTCCCACCGTCCGC  264

Qy    241 TACCAGCTCGAGGATGAGCACCGTGAGAAGTGCGTGTCGCGCTACAGCGCCACCAGGCTGT  300
          |||||||||||||||||| |||||| |||||| |||||||| ||||  |||||||||
Db    265 TACCAGCTCGAGGGTGAGTCACGAAAGTCATGTGAGCCGACACCGTGACCAGGCTCCCACCGTCCGC  324

Qy    301 GGTGCGGTGCGCTTCTGGTGTTCACTGCCAACAGCGACACGTGCTGTCGCGCCCCTA  360
          ||||  || ||| |||||||| |||||||||  |||||||||||| ||||||||||
Db    325 GCTCCGCGGCCTTCTGGTGTGTCACTGCCAACAGCGACACATGCGAGTTTTGTGCCCGTG  384

Qy    361 GAGTTGCGCGTCACAGCAGCCTCCGGCGCTCCGGATATCACCGTGTCATCCACATCAAT  420
          ||| |||||||||||||||||||| ||||  |||||||||  |||||||||||| |||||
Db    385 GAGCTGCAGGTGACGGAGGGGTCCGTTCTCCTCGCTATCACGGCGATCATCCATATCAAT  444

Qy    421 GAAGTAGTGCTCCTAGACGCCGGTTGGTGGCGGTTGGCTGCGACGAGAGGCGC  480
          |||||||||||| |||||||||||||||| |||| ||| ||||||||||||||
Db    445 GAAGTAGTGCTCCTGGGCGGGTTGGTGGCGGTTGGCTGCGCCAGAGAGGGGCAGC  504

Qy    481 CACGTAGTGTTGGCGTGGCTCTCCCGACCCCATGAGCGTCTCACATCCGCTAC  540
          |||||||| |||||| ||||||||||||||||  |||||||||||||||||
Db    505 CACGTTGGTTGTGGCGCTGCCACCTCTGGAGCACCTGACCACCACATCCGCGATAT  564

Qy    541 GAGGTGGACGTTCCGGCCGCCGGCAACCGGCAGGAGGGGTGGAGATCCTGGAG  600
          |||||||||| |||||| ||||||  ||||  |||||||| |||| ||||||||
Db    565 GAAGTGGACGTGTCGGCAGCCAACCGGCAGGAGGAGGACAACAAAGGGTGGAGGTCCTGGAA  624

Qy    601 GGCCCGACCGAGTGTGTGCTGAGCAACCTGGGGGGG  660
          |||||| |||||||||| |||||||||||||||
Db    625 GGCCCACCACTGAGTGTGTTCTGAGCAACCTCCGGGGAGACGTACACCTTCGCGTGTT  684

Qy    661 CGGCGCGTATGGCCGGTGAGCCGAGCTTCGGCGGCCCGTGG  720
          ||||||||| |||| |||| ||||||  ||||||||
Db    685 CGAGCGGCGCATGGCCGGATTCGAGGCCGTGCCTCGTCCGAGCCCGTG  744

Qy    721 TCGCTGCTGACGGCCTAGCGACCCCCTCGACCCCATCGTG  780
          ||||| ||||| ||||| ||||||||||  ||||||||||
Db    745 TCACTACTGACCGCCTAGCGACCCCTCTCATCTTGACGCTGTTCCATTCTGGTC  804

Qy    781 GTCATCCTGGTGCTGCTGAGGCGTGCTGCCGCGGGCGTCTGAAGCAG  840
          |||| ||||| ||||| |||| |||| ||||||||||||||||||||
Db    805 CTCATCTCGGCGGTTCGCTGAGACGGTTCGTGACCCTCCGTCTCCGCGGGACTCTGCAGCAG  864

Qy    841 AAGATCTGGCCGGCATCCCAGAGCGGAGTTGAAGGCTCTTCACCACCCAC  900
          |||| |||||||| |||||||||| ||||| |||||| ||| ||||||||
Db    865 AAGATCTGGCCGGCATCCCAAGCCACAGCAAGAGTTTGAGGGTCTTCTTCACCACCCAC  924

Qy    901 AAGGGTAACTTCAGCAGCTGGCCGTACCAGAATGATGGCGTGCCTGGTGGCAGCCTGC  960
          |||||||||||||||||| ||||| ||||||| |||| ||||||||||||||||||
Db    925 AAGGGTAACTTCAGCAGCTGGCCGTGATGGTTGTCTGTGGTGGTGAGCGGGGC  984

Qy    961 ACCCCTTCACGGAGGACCACCTGTTCCCAGGAAGTCCTTCCAGAGGCGCTGTGGGGG  1020
          | ||||||| ||||| ||||||||  |||||| ||||||||||||||||||||
Db    985 AGCTCCTTCCCTGAGGATCACCTTCCCACTAGAGGTCCTCCAGAGCCACACCGCTGGGCA  1044

Qy   1021 ACGATGCAGGCAGTGGAGCCGGGACAGATGATGAGGGCCCCCTGCTGAGCCAGTGGGGC  1080
          ||   |||||||| |||||||||| |||| |||||||| ||| ||||||||||||||
Db   1045 GTGACTCAGGCTGGGGGACCCAGGGACCAGGCAGATGATGAGGGGCCCCTTACTGACTGAGCCGGTGGGGC  1104
```

```
Db   1360 GCCTCAGAAGCATCCTTCTGCTCATCCTGCTTGGCCTGGCCCTGAGACCTGGGCCAGGAGGGGGCC  1419
          |||||||||||||||||||||||||||||||||||||||||||||| |||
Qy   1258 TCTGCTGCCAGCTTTGAGTAGTACCATATCCTGGACACCGAGCCGCTCCCAGCTCTTGCGTCCATGG  1317

Db   1420 TCAGCTGCCAGCTTTGAGTATATCTATCCTGGATCCGCACCCCACCCCAACCAGCTCCTGCGCCCAAGA  1479
          ||||||||||||||||||||| | || ||||||  ||| ||||||||||||
Qy   1318 ACACTGTGCCGCTGAGCTGCCCCTACCACCCCACCCCACCCCCACCCCAACTAAAGTAGTACCTTGTGGTA  1377

Db   1480 GCCCTGCCCGCTGAGCTGCCCGCTGAGCTCTAAAGTAGTACTTGTACCTTGTGGTG  1539
          |||||||||||||||||  |||| ||||||||||||| ||||||||||||||
Qy   1378 TCTGACTCTGGCATCTCAACTGACTACAGCTCAGGGGACTCCAGGGGACTCCAGGGGGC  1437

Db   1540 TCTGACTCTGGCATCTCAACTGACTACAGCTCAGGGGACTCCCAGGGGGGACCCAGGGGGC  1599
          |||||||||||||||||||||||||||||||||||  |||| |||||||| ||||
Qy   1438 TTATCCGATGGCCCCTACTCCAACCCTTATGAGAACAGCAGCCCTGTGTCCCAGCGCTGAGCCT  1497

Db   1600 TCATCCAGTGGCGCCCCACTCCAACCCTTATGAGAACAGCAGCCCTGTGTCCCAGCGCTGAGCCT  1659
          |||||| |||||||| |||||||||||||||||||||||||||||||||||||||||||
Qy   1498 CTGCCCCCAGCTATGTGGCTTGCTCTTAG  1527
          ||||| ||||||| ||| |||
Db   1660 TCACCCCCGAACTATGTGACCTGCTCATAG  1689
```

RESULT 13

```
MUSERPR
LOCUS          MUSERPR                1741 bp    mRNA         linear   ROD 12-JUN-1993
DEFINITION     Mouse erythropoietin receptor (EPO-R) mRNA, complete cds.
ACCESSION      J04843
VERSION        J04843.1  GI:193090
KEYWORDS       erythropoietin receptor; hematopoietic growth factor receptor;
               transmembrane protein.
SOURCE         Mus musculus (house mouse)
  ORGANISM     Mus musculus
               Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
               Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE      1  (bases 1 to 1741)
  AUTHORS      D'Andrea,A.D., Lodish,H.F. and Wong,G.G.
  TITLE        Expression cloning of the murine erythropoietin receptor
  JOURNAL      Cell 57 (2), 277-285 (1989)
  MEDLINE      89195238
   PUBMED      2539263
COMMENT        Original source text: Mouse erythroleukemia cell line MEL subclone
               745, cDNA to mRNA, clone 190.
               Draft entry for [1] kindly provided by A.D. D'Andrea, 28-APR-1989.
FEATURES             Location/Qualifiers
     source          1..1741
                     /organism="Mus musculus"
                     /mol_type="mRNA"
                     /db_xref="taxon:10090"
     CDS             28..1551
                     /note="erythropoietin receptor precursor"
                     /codon_start=1
                     /protein_id="AAA37571.1"
                     /db_xref="GI:309219"
                     /translation="MDKLRVPLWPRVGPLCLLLAGAAWAPSSLPDKFESKAALLAS
                     RGSEELLCFTQRLEDLVCFWEEAASSGMDFNYSFSYQLEGESRKSCSLHQAPTVRGSV
                     RFWCSLPTADTSSFVPLELQVTEASGSPRYHRIIHINEVVLLDAPAGLLARRAEEGSH
                     VVLRWLPPPGAPMTHIRYEVDVSAGNRAGGTQRVEVLEGRTECVLSNLRGGTRYTFA
                     VRARMARPSFSGFWSAWSEPASLLTASDLDPLILTLSLILVLJSLLLTVLALLSHRRT
                     LQQKIWPGIPSPESEFEGLFTTHKGNFQLWLLQRDGCLWWSPGSSFPBDPPAHLEVLS
                     EPRWAVTQAGDPGADDEGPLLEPVGSEHAQDYTYLVLDKWLLPRTPCSSNLSGPGGSVD
                     PVTMDBRASETSSCPSDLASKPRPEGTSPSSFEYTILDPSSQLLCPRALPPELPPTPPH
                     LKRYLYLVVSDSGISTDYSSGSGSQQVHGDSSDGPYSHPYENSLVPDSEPLHPGYVACS"
     sig_peptide     28..99
                     /note="erythropoietin receptor signal peptide"
     mat_peptide     100..1548
                     /product="erythropoietin receptor"
ORIGIN         271 bp upstream of XhoI site.

Query Match     70.8%;  Score 1080.6;  DB 10;  Length 1741;
Best Local Similarity   82.2%;  Pred. No. 2.2e-179;
Matches 1255;  Conservative  0;  Mismatches 269;  Indels  3;  Gaps  1;
```

```
              Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
              USA
REMARK        NIH-MGC Project URL: http://mgc.nci.nih.gov
COMMENT       Contact: MGC help desk
              Email: cgapbs-r@mail.nih.gov
              Tissue Procurement: Lothar Hennighausen Ph.D., Robin Humphreys
              cDNA Library Preparation: Life Technologies, Inc.
              cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
              DNA Sequencing by: Sequencing Group at the Stanford Human Genome
              Center, Stanford University School of Medicine, Stanford, CA 94305
              Web site:      http://www-shgc.stanford.edu
              Contact:  (Dickson, Mark) mcd@paxil.stanford.edu
              Dickson, M., Schmutz, J., Grimwood, J., Rodriguez, A., and Myers,
              R. M.

              Clone distribution: MGC clone distribution information can be found
              through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
              Series: IRAK Plate: 7 Row: i Column: 4
              This clone was selected for full length sequencing because it
              passed the following selection criteria: Hexamer frequency ORF
              analysis.

FEATURES          Location/Qualifiers
     source       1..1695
                  /organism="Mus musculus"
                  /mol_type="mRNA"
                  /strain="mix FVB/N, C57BL/6J"
                  /db_xref="taxon:10090"
                  /clone="IMAGE:3487672"
                  /tissue_type="Mammary tumor. WAP-TGF alpha model. 7 months
                  old, gross tissue."
                  /clone_lib="NCI_CGAP_Mam5"
                  /lab_host="DH10B"
                  /note="Vector: pCMV-SPORT6"
     gene         <1..1695
                  /gene="Epor"
                  /db_xref="LocusID:13857"
                  /db_xref="MGI:95408"
     CDS          <1..1522
                  /gene="Epor"
                  /codon_start=2
                  /product="Epor protein"
                  /protein_id="AAH03953.1"
                  /db_xref="GI:1327241"
                  /db_xref="LocusID:13857"
                  /db_xref="MGI:95408"
                  /translation="DKLRVPLWPRVGPLCLLLAGAAWAPSPLPDKFESKAALLASR
                  GSEELLCFTQRLEDLVCFWEEAASSGMDFNYSFSYQLEGSRKSCSLHQAPTVRGSVR
                  FWCSLPTADTSFVPLELQVTEASGSPRYHIRIHINEVVLLDAPAGLLARRAEEGSHV
                  VLRWLPPGAPMTHIRYEVDVSAGNRAGGTQRVEVLEGRTECVLSNLRGGTRYTFAV
                  RARMAEPSFGGFWSAWSEPASLLTASDLDPLILTLSLILVLISLLTVLALLSHRRTL
                  QQKIWGIPSPESEFEGLFTTHKGNFQLWLLQRDGCLWWSPGSSFPEDPAHLEVLSE
                  PRWAVTQAGDPGADDEGPLLEPVGSEHAQDTYLVLDKWLLPRTPCSENLSGPGGSVDP
                  VTMDEASETSSCPSDLASKRPEGTSPSSFEYTILDPSSQLLCPRALPPELPPTPPHL
                  KYLYLVVSDSGISTDYSSGSQGVHGDSSDGPYSHYENSLVPDSEPLHPGYVACS"
     misc_feature 431..727
                  /gene="Epor"
                  /note="FN3; Region: Fibronectin type 3 domain"
                  /db_xref="CDD:cd00063"

ORIGIN
        Query Match      70.6%;     Score 1078.6;  DB 10;  Length 1695;
        Best Local Similarity  82.2%;     Pred. No. 4.9e-179;
        Matches 1253;    Conservative   0;  Mismatches 269;   Indels    3;  Gaps    1;

Qy    3 GGACCACCTCGGGGCGTCCCTGGCCCTGCTGCTGGCCGGGGCTGCCTGG  62
        |||||||||||||||   |||||||||  |||||| |||||||  ||||||||||  |
Db    1 GGCAAACTCAGGGTGCCCTGCCCTCGGCTAGGCCCCCTAGGCCCCTCCTACTTGCTGG  60

Qy   63 GGCGCCTGGCGCGCCCCCGCCTAACCTCCCGGACACCCAAGTTCGAGAGCAAAGCGGCCTT  122
        ||||||||||||||  ||||||  |||||||  |||||||||  |||||||||||||||||
Db   61 GGCAGCCTGGGCACCTTCACCCCGCCTCCGGACCCCCAAGTTTGAGAGCAAAGCGGCCCT  120

Qy  123 GCTGGCGGCGGCCCGGGGCGCCGAAGAGTTCTGTTGCTTCACCGAGAGCGGTTGGAGGACTTGGT  182
```

```
Qy  1081 AGTGAGCAATGCCCAGGATACCTATCTGGTGCTGGACAAATGGTTGCTGCCCCGGAACCCG  1140
         ||||||||||||||||||||  ||||  |||||||  |||||||||||||||||||||||||
Db  1105 AGTGAGCAATGCCCAGGAGACCACCTACTTGGTATTGGATAAGTGGTTGCTGCCCCGGACCCCA  1164

Qy  1141 CCCAGTGAGGACCTCCCAGGGCCTGGTGCGCAGTGTGACATAGTGGCCATGGATGAGGC  1200
         |||||||||||||||||||||  ||||||||  |||||||  |||  |||||||||||||
Db  1165 TGCAGTGAGAACCTCTCAGGGCCTGGGCCAGTCTGGACCCGTGGACCTGTGACTATGGATGAAGCT  1224

Qy  1201 TCAGAAGACATCCTCCTGCTCCATCGTGCTTGGCCTCGAAGCCCAGAGGGAGCCTCT  1260
         |||  |||  |||  |||  |||||||||  ||||||||||||||||||  |||||||
Db  1225 TCAGAAACATCTTCCTGCCCGTCTCGACTTGGCCTCAAAGCCCAGGCCAGGGCACCTCA  1284

Qy  1261 GCTGCCAGCTTTGAGTACACTATCCTGGACCCCAGCTCCCAGCTCTTGCGTCCATGGACA  1320
         |||||||||||||||||||  ||||||||||||||||||  |||||  |||  ||||||||
Db  1285 CCTTCCAGCTTTGAGTACACCACCATCCTGGACCCCCAGCTCTCAGCTCCTGTGCCTCGGGGCA  1344

Qy  1321 CTGTGCCCTGAGCTGCCCCCTACCCCACCCCACCCCTAAGTACCTGTACCTTGTGGTATCT  1380
         |||||||||||||||  |||||||||||||||  |||||||||||||||||||  ||||||
Db  1345 CTGCCTCCCGAGCTACCTCCCACTCCCACCCCACCTACTTGAAGTACCTATACCTTGTGTGTCC  1404

Qy  1381 GACTCTGGGGATCTCAACTGACTACAGCTCCAGGGGAGCTCCCAGGGAGCCCAAGGGGCTTA  1440
         |||||||||||||||||||  |||  ||||||||||  ||||||  |||||||||||||||
Db  1405 GATTCTGGGGATCTCAACAGATTACAGTTCGGGGGTCTTCAGGGAGTCCACGGGGACTCA  1464

Qy  1441 TCCGATGGCCCCTACTCCAACCTTATGAGAACAGCAGCCTTATCCCAGCCCGTGAGCCTCTG  1500
         |||||||||||||||||||  |||||||||||||  |||||||||  |||||||||||||||
Db  1465 TCTGATGGCCCCTACTCCCACCCCCATGAGAACAGCAGCCTTGTCCCAGACTCAGAGCCCTCTG  1524

Qy  1501 CCCCCCAGCTCTATGTGGCTTGCTTCTTAG  1527

Db  1525 CATCCCGGCTATGTGGCCTGCTCCTAG  1551
```

```
RESULT 14
BC003953
LOCUS       BC003953               1695 bp    mRNA      linear   ROD 13-FEB-2004
DEFINITION  Mus musculus erythropoietin receptor, mRNA (cDNA clone
            IMAGE:3487672), partial cds.
ACCESSION   BC003953
VERSION     BC003953.1  GI:13278240
KEYWORDS    .
SOURCE      Mus musculus (house mouse)
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 1695)
  AUTHORS   Strausberg,R.L., Feingold,E.A., Grouse,L.H., Derge,J.G.,
            Klausner,R.D., Collins,F.S., Wagner,L., Shenmen,C.M., Schuler,G.D.,
            Altschul,S.F., Zeeberg,B., Buetow,K.H., Schaefer,C.F., Bhat,N.K.,
            Hopkins,R.F., Jordan,H., Moore,T., Max,S.I., Wang,J., Hsieh,F.,
            Diatchenko,L., Marusina,K., Farmer,A.A., Rubin,G.M., Hong,L.,
            Stapleton,M., Soares,M.B., Bonaldo,M.F., Casavant,T.L.,
            Scheetz,T.E., Brownstein,M.J., Usdin,T.B., Toshiyuki,S.,
            Carninci,P., Prange,C., Raha,S.S., Loquellano,N.A., Peters,G.J.,
            Abramson,R.D., Mullahy,S.J., Bosak,S.A., McEwan,P.J.,
            McKernan,K.J., Malek,J.A., Gunaratne,P.H., Richards,S.,
            Worley,K.C., Hale,S., Garcia,A.M., Gay,L.J., Hulyk,S.W.,
            Villalon,D.K., Muzny,D.M., Sodergren,E.J., Lu,X., Gibbs,R.A.,
            Fahey,J., Helton,E., Ketteman,M., Madan,A., Rodrigues,S.,
            Sanchez,A., Whiting,M., Madan,A., Young,A.C., Shevchenko,Y.,
            Bouffard,G.G., Blakesley,R.W., Touchman,J.W., Green,E.D.,
            Dickson,M.C., Rodriguez,A.C., Grimwood,J., Schmutz,J., Myers,R.M.,
            Butterfield,Y.S., Krzywinski,M.I., Skalska,U., Smailus,D.E.,
            Schnerch,A., Schein,J.E., Jones,S.J. and Marra,M.A.
  TITLE     Generation and initial analysis of more than 15,000 full-length
            human and mouse cDNA sequences
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
   PUBMED   12477932
REFERENCE   2  (bases 1 to 1695)
  AUTHORS   Strausberg,R.
  TITLE     Direct Submission
  JOURNAL   Submitted (28-FEB-2001) National Institutes of Health, Mammalian
            Gene Collection (MGC), Cancer Genomics Office, National Cancer
```

Db   1198 AGAAACATCTTCCTGCCCGTCTGACTTGGCCTGGCCTCAAAGCCCAGGGCCACCTCACC 1257
            |||||| ||||||| ||||||| |||||| |||||||||
Qy   1263 TGCCAGCTTGAGTACACTACCATCCTGGACCCCAGTCCTTCTGCCTCTCCATGGACACT 1322

Db   1258 TTCCAGCTTGAGTACACCATCCTGGACCCCAGCTCTCAGCTTCCTGTGCCCTGGGCACT 1317
            |||||||||||||||| |||||||||||||||||| |||||||||||||
Qy   1323 GTGCCCTGAGTGCCCCTACCCCACCCACCTAAAGTACTTGTGGTATCGA 1382

Db   1318 GCCTCCCGAGCTACCTCCACCTCCACTCACTTGAAGTACCTATACCTTGTGTGTCCGA 1377
            |||||| ||| |||||||| |||||| |||||||||||||||
Qy   1383 CTCTGGCATCTCAACTGACTACACAGCTCAGGGGACTCCCAGGGAGCCCAAGGGGCTTATC 1442

Db   1378 TTCTGGCATCTCAACAGATTACAGTTCGGGGGGCTCTCAGGGAGTCCACGGGGACTCATC 1437
            ||||||||| ||||||||||||||| |||||||
Qy   1443 CGATGGCCCCTACTCCAACCCTTATGAGAACAGCCTTATCCCAGCCCGCTGAGCCTCTGCC 1502

Db   1438 TGATGGCCCCTACTCCCACCCCCATGAGAACAGCCTTGTCCCAGACTCAGAGCCTCTGCA 1497
            |||||||| ||||||| |||||| |||||||||||||
Qy   1503 CCCCAGCTATGTGGCTTGCTCTTAG 1527

Db   1498 TCCCGGCTATGTGGCCCTGCTCCTAG 1522

RESULT 15
BC046282
LOCUS   BC046282   1769 bp   mRNA   linear   ROD 07-OCT-2003
DEFINITION   Mus musculus erythropoietin receptor, mRNA (cDNA clone MGC:54669
               IMAGE:6438591), complete cds.
ACCESSION   BC046282
VERSION   BC046282.1 GI:28279360
KEYWORDS   MGC.
SOURCE   Mus musculus (house mouse)
  ORGANISM   Mus musculus
               Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
               Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1 (bases 1 to 1769)
  AUTHORS   Strausberg,R.L., Feingold,E.A., Grouse,L.H., Derge,J.G.,
               Klausner,R.D., Collins,F.S., Wagner,L., Shenmen,C.M., Schuler,G.D.,
               Altschul,S.F., Zeeberg,B., Buetow,K.H., Schaefer,C.F., Bhat,N.K.,
               Hopkins,R.F., Jordan,H., Moore,T., Max,S.I., Wang,J., Hsieh,F.,
               Diatchenko,L., Marusina,K., Farmer,A.A., Rubin,G.M., Hong,L.,
               Stapleton,M., Soares,M.B., Bonaldo,M.F., Casavant,T.L.,
               Scheetz,T.E., Brownstein,M.J., Usdin,T.B., Toshiyuki,S.,
               Carninci,P., Prange,C., Raha,S.S., Loquellano,N.A., Peters,G.J.,
               Abramson,R.D., Mullahy,S.J., Bosak,S.A., McEwan,P.J.,
               McKernan,K.J., Malek,J.A., Gunaratne,P.H., Richards,S.,
               Worley,K.C., Hale,S., Garcia,A.M., Gay,L.J., Hulyk,S.W.,
               Villalon,D.K., Muzny,D.M., Sodergren,E.J., Lu,X., Gibbs,R.A.,
               Fahey,J., Helton,E., Ketteman,M., Madan,A., Rodrigues,S.,
               Sanchez,A., Whiting,M., Madan,A., Young,A.C., Shevchenko,Y.,
               Bouffard,G.G., Blakesley,R.W., Touchman,J.W., Green,E.D.,
               Dickson,M.C., Rodriguez,A.C., Grimwood,J., Schmutz,J., Myers,R.M.,
               Butterfield,Y.S., Krzywinski,M.I., Skalska,U., Smailus,D.E.,
               Schnerch,A., Schein,J.E., Jones,S.J. and Marra,M.A.
  TITLE   Generation and initial analysis of more than 15,000 full-length
               human and mouse cDNA sequences
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
  MEDLINE   22388257
   PUBMED   12477932
REFERENCE   2 (bases 1 to 1769)
  AUTHORS   Strausberg,R.
  TITLE   Direct Submission
  JOURNAL   Submitted (31-JAN-2003) National Institutes of Health, Mammalian
               Gene Collection (MGC), Cancer Genomics Office, National Cancer
               Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
               USA
REMARK   NIH-MGC Project URL: http://mgc.nci.nih.gov
COMMENT   Contact: MGC help desk
               Email: cgapbs-r@mail.nih.gov
               Tissue Procurement: Gilbert Smith, Ph.D.
               cDNA Library Preparation: Life Technologies, Inc.
               cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
               DNA Sequencing by: National Institutes of Health Intramural

Db   121   GCTGGCATCCCGGGCTCCGAGAACTTCTGTGCTTCACCCACGCGCTTGGAGACTTGGT 180
            |||||| |||||| ||||||| |||||| ||||||
Qy   183   GTGTTTCTCGGGAGGAAGCGGCGAGCGGTGGGGCCCGGCCAACTACAGCTTCTCCTA 242

Db   181   GTGTTTTCTGGGAGGAAGCGGCGAGCTCCGGGATGGAC---TTCAACTACAGTTCTCATA 237
            ||||||||||||||| ||||||| ||||||
Qy   243   CCAGCTCGAGGATGAGCGAGCCATGGAAGCTGTGTCGCCTCGACAGGCTCCCAGGCTTCGTGG 302

Db   238   CCAGCTCGAGGGTGAGTCAGAAAGTCATGTAGCTGTGCACCAGGCTCCCACGGCTCCGCGG 297
            |||||||||||||| ||||||||||||| ||||||
Qy   303   TGGGGTGGCGCTTCGTGTTCCTGCTACAGCCGACACGTCGAGCTTCGTGTCCCCTAGA 362

Db   298   CTCCGTGCGTTTCTGGTGTTCACTGCCAACAGCGGACACATCGAGTTTTGTCCGCTGGA 357
            |||||| |||||| ||||||||||||| ||||||
Qy   363   GTTGCGGTCACAGCAGCCTCCGGCCGCTCCGGATATCACCGTGTCATCCACATCAATGA 422

Db   358   GCTGCAGGTGACGGAGGCGTCCGGTTCCTCGCTATCACCGCATCATCCATATCAATGA 417
            ||||||||||||| |||||| |||||||
Qy   423   AGTAGTGCTCCTAGACGCCCCCTGTGGGCTGCTGGTGTGTGACGAGAGGCGGCCA 482

Db   418   AGTAGTGCTCCTAGACGCCCCCTGGCTCCTGGTGCGGCTGCTGGGCGCAGAGAGGGCAGCCA 477
            |||||||||||||| ||||||||||||| ||||||
Qy   483   CGTAGTGTTGCGCGTCTGGCTGGTCCCCGCCTGAGACACCCATGAGCGTCTCACATCCGCTACGA 542

Db   478   CGTGGTGTGCGCGTCGTGCTGCTGGTGCTGCTGCCACCTTCCTGAGCACCTGATGACCACACCCACATCCGAGGG 537
            |||||| ||||||||||||| |||||||||||| |||||||
Qy   543   GGTGACGTCTCGGCCCGGCGCAGCAACGGCGGCAGGGAGCCGTACAGAGAGGGTGGAGATTCCTGGAGGG 602

Db   538   AGTGGACGTGTGGCAGGCAACGGCCAGCAACCGGGAGGAGCGAACAAAGGGTGGAGGTCCTGGAAGG 597
            |||||||||||||| ||||||||||||| ||||||
Qy   603   CCGCACCGAGTGTGTGCTGAGCAACTGCTGCGAGCAACCTGGTGCCGACGGCGCGCCCACCTTCGCCCTCCG 662

Db   598   CCGCACTGAGTGTGTTCTGAGCAACTGCTGCGAGCAACCTGGTGCCGACGGCGCGCCTACACCTTCGCCGTTCG 657
            |||||||||||||| ||||||||||||||| ||||||
Qy   663   CGCGCGTATGGCTGAGCGGCGAGCTTCGGCGACGGCTTCGTGGAGCCTGTGTC 722

Db   658   AGCGCGCGCCATGGCTGAGCGGCCGAGCAGCTTCAGCGGATTCTGAGTGCCTGGAGTCCTGAGCCCCGCGTC 717
            |||||| |||||| ||||||||||||| ||||||
Qy   723   GCTGCTGACGGCCTAGCGACCTGGACCCCTCATCGTCTCCCTGACCGTCTTCTCATTCTGTGTCCT 782

Db   718   ACTACTGACGGCCTAGGACCGCTGGACCCCTTCTCATCTTGACCGTGTTCATCTCTGTCCTGGAAGG 777
            |||||| ||||||||||||| |||||||||||| ||||||
Qy   783   CATCCTGGTCGTGTGCTCTGACCGTGTCTCGCGCTCTCCCCACCGGGCCTCTGAAGCAGAGA 842

Db   778   CATCTCGTGTTGCTGAGCGTTCTGACGGTTCTGGCCCGTGTCCCACCGCCGGGACTCTGCAGCAGAGA 837
            |||||||||||| ||||||||||||||| ||||||
Qy   843   GATCTGGCCTGGCATCCGAGCCCAGAGAGCGAGTTGAAGGCCTTCTTCACCACCCACAA 902

Db   838   GATCTGGCCTGGCATCCCAAGCCCAGAGAGCGAGTTGAGGGTCTTCTTCACCACCACCACAA 897
            |||||| ||||||||||||| |||||||||||||| ||||||
Qy   903   GGGTAACTTCCAGCTGTGGCTGTACCAGGCTGCCTGTGTGGTGGAGCCCCTGCACGGGGGGAC 962

Db   898   GGGTAACTTCCAGCTGTGGCTGCGCTGCAGCGTGATGCTGTTGTCTGTGTGTGGAGCCCGCGGCAG 957
            |||||||||||||| ||||||||||||| ||||||
Qy   963   CCCCTTCACGGAGGACGAGGACCACCTGCTTCCCTGCAGTGGGCAGCAGTGGCGCAG 1022

Db   958   CTCCTTCCCTGAGGATCCACCTGCCCACCTCCCTCAGAGAGTGCCTGCTTACTGGAGCCAGTGGGCAGT 1017
            |||||| |||||||||||| ||||||| ||||||
Qy   1023 GATGCAGGCAGTGGGACCGGCGGACAGATGATGAGGGCCGCCTGCTGGGCCAGTGGGCAG 1082

Db   1018 GACTCAGGCTGGGGACCACCCAGGGGCAGATGATGAGGGCCCCCTACTGGAGCCGGGCAG 1077
            |||||| |||||||||||| ||||||||||||| ||||||
Qy   1083 TGAGCATGCCCAGGATACCTTATCTGGTCTGGACAAATGGTTGCTGCCCCGGAACCCGCC 1142

Db   1078 TGAGCATGCCCAGGATACCTTACTTGGTATTGGATAAGTGGTGGTTGCTGCCCCCCGGACCCCATG 1137
            |||||| ||||||||||||| ||||||||||||| ||||||
Qy   1143 CAGTGAGGGACCTCCCAGGGGCTGGTTGGCAGTGTGGCACATAGTGGACCATGGATGAAGGCTC 1202

Db   1138 CAGTGAGAACCTCTCAGGGCCTGGGGGCAGTGTGGCAGTCCCTGTGACCCCTGTGACTATTGGATGAAGAAGCTTC 1197
            |||||| ||||||||||||| ||||||| ||||||
Qy   1203 AGAAGCATCCTCCGTCATCTGCTTTGGCCTCAGGGCCCAGCCCAGAGGGAGCCCTGC 1262

Sequencing Center (NISC),
Gaithersburg, Maryland;
Web site: http://www.nisc.nih.gov/
Contact: nisc_mgc@nhgri.nih.gov
Akter,N., Ayele,K., Beckstrom-Sternberg,S.M., Benjamin,B.,
Blakesley,R.W., Bouffard,G.G., Breen,K., Brinkley,C., Brooks,S.,
Dietrich,N.L., Granite,S., Guan,X., Gupta,J., Haghighi,P.,
Hansen,N., Ho,S.-L., Karlins,E., Kwong,P., Laric,P., Legaspi,R.,
Maduro,Q.L., Masiello,C., Maskeri,B., Mastrian,S.D., McCloskey,J.C.,
McDowell,J., Pearson,R., Stantripop,S., Thomas,P.J., Touchman,J.W.,
Tsurgeon,C., Vogt,J.L., Walker,M.A., Wetherby,K.D., Wiggins,L.,
Young,A., Zhang,L.-H. and Green,E.D.

Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
Series: IRAK Plate: 100 Row: i Column: 3
This clone was selected for full length sequencing because it
passed the following selection criteria: matched mRNA gi: 6753765.

```
FEATURES             Location/Qualifiers
     source          1..1769
                     /organism="Mus musculus"
                     /mol_type="mRNA"
                     /strain="FVB/N-3"
                     /db_xref="taxon:10090"
                     /clone="MGC:54669 IMAGE:6438591"
                     /tissue_type="Mammary tumor. MMTV-LTR/INT3 model. 5 month
                     old mouse. Taken by biopsy."
                     /clone_lib="NCI_CGAP_Mam2"
                     /lab_host="DH10B"
                     /note="Vector: pCMV-SPORT6"
     gene            1..1769
                     /gene="Epor"
                     /db_xref="LocusID:13857"
                     /db_xref="MGI:95408"
     CDS             76..1599
                     /gene="Epor"
                     /codon_start=1
                     /product="Epor protein"
                     /protein_id="AAH46282.1"
                     /db_xref="GI:28279361"
                     /db_xref="LocusID:13857"
                     /db_xref="MGI:95408"
                     /translation="MDKLRVPLWPRVGPLCLLLAGAAWAPSPSLPDPKFESKAALLAS
                     RGSEELLCFTQRLEDLVCFWEEAASSGMDPNYSFSYQLEGESRKSCSLHQAPTVRGSV
                     RFWCSLPTADTSSFVPLELQVTEASGSPRYHRIIHINEVVLLDAPAGLLARRAEGSH
                     VVLRWLPPPGAPMTHIRYEVDVSAGNRAGGTQRVEVLEGRTECVLSNLRGGTRYTFA
                     VRARMAEPSFSGFWSAWSEPASLITASDLDPLIITLSLILIVLISLLLTVLALLSHRRT
                     LQQKIWPGIPSPESSFEGLFTTHKGNFQLWLLQRDGCLWWSPGSSFPEDPPAHIEVLS
                     EPRRAVTQAGDPGADDEGPLLEPVGSEHAQDTYLVLDKWLLPRTPCSSENLSGPGGSVD
                     PVTMDEASETSSCPSDLASKRPEGTSPSSFEYTILDFSSQLLCPRALPPELPPTPPH
                     LKYLYLVVSDSGISTDYSSGSQGVHGDSSDGPYSHPYENSLVPDSEPLHPGYVACS"
     misc_feature    508..741
                     /gene="Epor"
                     /note="FN3; Region: Fibronectin type 3 domain"
                     /db_xref="CDD:smart00060"
ORIGIN

Query Match         70.5%;    Score 1075.8;   DB 10;   Length 1769;
Best Local Similarity 82.0%;  Pred. No. 1.5e-178;
Matches 1252;  Conservative     0;  Mismatches  272;  Indels    3;  Gaps    1;
```

```
Qy     1 ATGGACCACCTCGGGGCGTCCCTCTGGCCCCGAGGTCGGCCCCTTTGTCTCCTGCTGGCT   60
         ||||||    || |||| || ||||||||||   ||| || || ||| ||||| ||| ||
Db    76 ATGGACAAACTCAGAGTGCCCCTCTGGCCCTCGGGTAGGCCCCTCCGGGCTCCTACTTGCT  135

Qy    61 GGGGGCCGCCTGGGCCCCCAGCCCTCCCTAACCTCCCCGGACCCCCAAGTTCGAGAGCAAGGCGGCC  120
         |||||    |||| ||| |||||| |||| || |||||||| ||||||||||||| ||||| ||||||
Db   136 GGGGGCACCTGGGCGCCCAGCCCTCACCCAGTGTCCTCCGGACCCCCAAGTTTGAGAGCAAAGCGGCC  195

Qy   121 TTGCTGGCGGCCCCGGGGGCCGAAGAGCTTCTGTGCTTCACCCAGCGGCTTGGAGGACTTG  180
         |||||||     ||| |||||| ||||||||||| |||| ||| |  ||| ||||||||||
Db   196 CTGCTGGCAGCCCCGGGGGCTCCGAAGAACTTCTGTGCTGTGTCCGGCTCCCAACGCTTGGAAGACTTG  255
```

	Qy/Db	position
Qy	GTGTGTTTCTGGGAGGAAGCGGCGAGCGGCTGGGGTGGCGAGGTGGGTGGC	181 – 240
Db	GTGTGTTTCTGGGAGGAAGCGGCGAGCTCCGGGATGGAC---TTCAACTACAGCTTCTCA	256 – 312
Qy	TACCAGTTGAGGATGAGCCATGGAGCTGTCGCTGGAGCTCCCACGGCTCGT	241 – 300
Db	TACCAGCTCGAGGGTGAGTCACGAAAGTCATGTAGCCTGCACCAGGTCCACCGTCCGC	313 – 372
Qy	GGTGCGGTGCGCTTCTGGTGTTCGCTCCTACAGCCGACACGTCGAGCTTCGTGCCCCTA	301 – 360
Db	GGCTCCGTGCGGTTTCTGGTGTTCACTGCCAACAGCGGACACATCGAGTTTGTGCCGCTG	373 – 432
Qy	GAGTTGCGCGTCACAGCAGCCTCCGGCGCTCCGGCGGATATCACCGTGTCATCCACATCAAT	361 – 420
Db	GAGCTGCAGGTGACGGAGGCGGTCCGGGTTCCTCGTCCGGCTCATCACCGGCATCATCCATATCAAT	433 – 492
Qy	GAAGTAGTGCTCCTAGACGCCCCGTGGGGCTTGGTGGCGCGGTTGGCTGACGAGAGCGGC	421 – 480
Db	GAAGTAGTGCTTCCTGGACCAACTTCCTGAGCACCTATGACCCACCCACATCCGATAT	493 – 552
Qy	CACGTAGTGTTGCGCTGGCTCCCGCCGCCCGCCCATGACACCCATGAGCGTCTCACATCCGCTAC	481 – 540
Db	CACGTGTGTGCTGCGCTGGCTGCCACCTCCTGAGCACCTATGACCCACCACCATCCGATAT	553 – 612
Qy	GAGGTGGACGTTCCGGCCGGCAACGCGCAGGGAGGCGTACAGAGGGTGGAGATCCTGGAG	541 – 600
Db	GAAGTGGACGTGTCGGCGGGCAACGCGGGACAGAGGGACAAAAGGGTGGAGGTGCCTGGAA	613 – 672
Qy	GGGCCGCACCGAGTGTGTGCTGAGCAACCTGCGGGGCCGGACCCGCTACACCTTCGCCGTC	601 – 660
Db	GGCCGCACTGAGTGTGTTCTGAGCAACCTGCGGGGACGGACCCGCTACACCTTCGCTGTT	673 – 732
Qy	CGCGCGCGTATGGCTGAGCCGAGCTTCGGCGGCCTTCGGAGCCCGTGTG	661 – 720
Db	CGAGCCCGCATGGCCGAGCCGAGCTTCAGCGGATTCTGGAGTGCTGGTCTGAGCCCCCG	733 – 792
Qy	TCGCTGCTGAGCGCTAGCGACCTGGACCCCCTCATCCTGACGGCTCTCCCTGTG	721 – 780
Db	TCACTACTGACCGCTAGCGACCTGGACCCTCTCATCTTGACCGTCTCATTCTGTC	793 – 852
Qy	GTCATCCTGGTGTCTGCTGCTGACGGTTCTGGCCCTGCGTGTCTGCAGCAG	781 – 840
Db	CTCATCTCGTTGTTGCTGACGGTTCTGGCCCTGCTGTCTCGTGAAGCAG	853 – 912
Qy	AAGATCTGGCCTGGCATCCCAGAGAGCGAGTTGAAGGCCTCTTCCACCACCAC	841 – 900
Db	AAGATCTGGCCTGGCATCCCAAGCCCAGAGAGCGAGTTTGAGGGTCTTCTTCACCACCAC	913 – 972
Qy	AAGGGTAACTTCCAGTGTGCTGTACCAGAATGATGGCTCCCTGGTGGAGCCCCTGC	901 – 960
Db	AAGGGTAACTTCCAGCTGTGGCTGCTGCAGCGCGATGGTGTCTGTGTGGGAGCCCCGGGC	973 – 1032
Qy	ACCCCCTTCACGGAGGCAGTGGAGCCCGGGACAGATGAGGGGCCCCGTGCTGCTGGGG	961 – 1020
Db	AGCTCCTTCCCTGAAGATCCACCTGCCCACCTAGAGGTCCTCTCAGAGCCCACGTCGGGCA	1033 – 1092
Qy	ACGATGGAGGCAGTGGAGCCCGGGACAGAGATGATGAGGGGCCCCAGTGGGC	1021 – 1080
Db	GTTACTCAGCTGGGGGACCCGGGGGCAGAGGGGCCCTTACTGGAGCCGGTGGGC	1093 – 1152
Qy	AGTGAGGCATGCCAGATACCTATCTGGTGCTGGACAAATGGTTGCTGCCCGGAACCCG	1081 – 1140
Db	AGTGAGCATGCCCAGGACACCTACTTGGTATTGGATAAGTGGTTGCTGCCCCGGACCCCA	1153 – 1212
Qy	CCCAGTGAGGACCTCCGGACCTAACCTCCCGGACCCCAAGTTCGAGAGGACAAGCGGCC	1141 – 1200
Db	TGCAGTGAGAACCTCTCAGGGCCTGGGGGCAGTGTGGACCCTGTGACTATGGATGGAGGCT	1213 – 1272
Qy	TCAGAAGCATCCTCCTGTCTCCGGAAGAGCTTCTGTGCTTCACCCAGCGGCGGTTGGAGGACTTG	1201 – 1260
Db	TCAGAAACATCTTCCTGCCCGTCTCCAAGCCCAGGCCCAGAGGGCACCTCA	1273 – 1332
Qy	GCTGCCAGCTTTGAGTACACTATCCTGGACCCCAGCTCCCAGCTCTTGCCATGGACA	1261 – 1320

```
Db          |||||||||||||||||| |||||||||||||||  |||||||  ||||| | ||| || ||
      1333  CCTTCCAGCTTTGAGTACACCATCCTGGACCCCAGCTCTCAGCTCCTGTGCCCTCGGGCA 1392

Qy    1321  CTGTGCCCTGAGCTGCCCCCTACCCCACCCCACCTAAAGTACCTGTACCTTGTGGTATCT 1380
            |||  || ||||| |||| |||| ||||  ||| |||||||| ||||||||||  |||| |
Db    1393  CTGCCTCCCGAGCTACCTCCCACTCCACCTCACTTGAAGTACCTATACCTTGTGGTGTCC 1452

Qy    1381  GACTCTGGCATCTCAACTGACTACAGCTCAGGGGACTCCCAGGGAGCCCAAGGGGGCTTA 1440
            |||||||||||| |||||||| |||| ||| ||| ||||||| ||| |||| || ||
Db    1453  GATTCTGGCATCTCAACAGATTACAGTTCGGGGGGCTCTCAGGGAGTCCACGGGGACTCA 1512

Qy    1441  TCCGATGGCCCCTACTCCAACCCTTATGAGAACAGCCTTATCCCAGCCGCGCTGAGCCTCTG 1500
            ||||||||||||| |||||||| ||||||||||||||| ||||||||| || ||||||||
Db    1513  TCTGATGGCCCCTACTCCCACCCCTATGAGAACAGCCTTGTCCCAGACTCAGAGCCTCTG 1572

Qy    1501  CCCCCCAGCTATGTGGCTTGCTCTTAG 1527
            | ||| |||||||||||| |||||| ||
Db    1573  CATCCCGGCTATGTGGCCTGCTCCTAG 1599
```

Search completed: April 21, 2005, 01:13:11
Job time : 6675 secs